Measuring Attitudes toward Public Spending Using a Multivariate Tax Summary Experiment ⬛ ⬢

Lucy Barnes, Jack Blumenau, and Benjamin E. Lauderdale
University College London

Abstract: *It is difficult to measure public views on trade-offs between spending priorities because public understanding of existing government spending is limited and the budgetary problem is complicated. We present a new measurement strategy using a continuous treatment, multivariate choice experiment. The experiment proposes deficit-neutral bundles of changes in spending and taxation, allowing us to investigate attitudes toward modifications to the existing budget. We then use a structural choice model to estimate public preferences over spending categories and the taxation level, on average and as a function of respondent attributes. In our application, we find that the UK public favors paying more in tax to finance large spending increases across major budget categories, that spending preferences are multidimensional, and that younger people prefer lower levels of taxation and spending than older people.*

How much money should the government collect in taxes, and how should it spend these resources across different policy areas? These two questions about the public budget are among the most important choices that governments make: They both change the material positions of citizens and enact political and social priorities. With this in mind, pollsters and political scientists have long sought to measure public attitudes regarding how much the government should tax and spend overall, and how much citizens want spent on different things. But despite the importance of taxation and spending itself, and of public opinion about taxation and spending to democratic governance, we lack good strategies for measuring these basic political preferences.

The fundamental tension in understanding and measuring public opinion on government budget choices is the collision of the intrinsic complexity of budgeting with the well-known limitations of citizens' understanding and engagement with public policy. How much information can we reasonably expect to extract about the public's views regarding taxation and government spending when very few members of the public have previously considered what their ideal budget might look like, let alone tried to articulate it? Existing studies tend either to ask simple questions, which make life easy for respondents, but avoid the complex realities of budgeting; or to maintain a high degree of fidelity to the budget process at the cost of very demanding response tasks for respondents. In either case, we may lack confidence that these articulated preferences are what will be relevant to public responses to the policies that governments in fact enact.

This article proposes a strategy for inferring a distribution of coherent public budget preferences from accessible questions about trade-offs over subsets of budget categories. We demonstrate this approach using a new

Lucy Barnes is Associate Professor in Comparative Politics, Department of Political Science, University College London, 29-31 Tavistock Square, London WC1H 9QU, United Kingdom (l.barnes@ucl.ac.uk). Jack Blumenau is Lecturer in Political Science and Quantitative Research Methods, Department of Political Science, University College London, 29-31 Tavistock Square, London WC1H 9QU, United Kingdom (j.blumenau@ucl.ac.uk). Benjamin E. Lauderdale is Professor of Political Science, Department of Political Science, University College London, 29-31 Tavistock Square, London WC1H 9QU, United Kingdom (b.lauderdale@ucl.ac.uk).

[Correction added on 12 October 2021, after first online publication: The copyright line was changed.]

survey experiment of UK citizens.[1] Every UK taxpayer receives a tax summary at the end of the financial year that indicates how individuals' income tax and national insurance contributions are proportionally allocated across a variety of public spending categories. Using these as the baseline for status quo levels of taxation and spending, respondents to our experiment are asked whether they prefer the status quo spending or randomly generated, deficit-neutral proposed bundles of changes in tax levels and spending categories. Our experimental design ensures that survey respondents have only to consider the relative merits of the status quo versus a proposal that is identical to the status quo on all but three potential spending areas and, possibly, the tax level. This simplifies the choice task faced by respondents (versus a full budget allocation), but it still maintains the essential quantitative budget constraint that spending changes and tax changes add up.[2]

The limited data we collect from each individual do not contain sufficient information to estimate preferences at the individual level. Instead, we use a structural choice model to translate the choices from our survey into estimates of the distribution of preferences across the UK public. We also estimate how average preferences vary as a function of demographic and political attributes of citizens. Our general strategy is to provide citizens with a set of questions that they could answer if they had well-formed budget preferences, observe the choices they make, and translate these choices back into the distribution of well-formed preferences across citizens who would have generated those responses. We acknowledge that most respondents may not in fact have well-formed budget preferences. One motivation for our design is that if citizens were asked to articulate such a budget, starting from the status quo and contemplating the trade-offs involved in changing it is a plausible strategy. Another motivation is that real budgets must add up, and so it is worth knowing the coherent budget that is most consistent with the spending and tax trade-offs that citizens are willing to endorse, even if individual citizens could not articulate a full budget themselves.

We identify three major substantive findings. First, on average, UK citizens express a willingness to support higher taxes in return for increased expenditure across a range of spending categories, which aggregate to an overall increase of 7% of current levels. Second, preferences over spending areas are multidimensional; it is not simply the case that some people want more spending and others want less or that there are "left-wing" budget areas and "right-wing" budget areas. Third, 18–34-year-olds want lower tax and spending levels than those older than 55, both unconditionally and holding constant their voting preferences, degree status, and income. This is notable given the very large age gradient in voting in recent elections, which have seen much higher support for the Labour Party among young people than older people. In contrast to much political commentary, young people in the UK are not left-wing in the sense of wanting a substantially larger government. We report both an internal validation, using our model to predict responses to held-out survey responses from our experiment, and an external validation, using our model to predict responses to novel budgetary proposals in a new preregistered survey.

Our study contributes to several literatures in political science. First, knowing what policies people support is necessary for answering questions about elite responsiveness to public preferences (Lax and Phillips 2009b). Much of this literature does not consider the responsiveness of government spending, focusing instead on issues where spending is a second-order concern: gay rights (Lax and Phillips 2009a), abortion (Arceneaux 2002), and support for judicial nominees (Kastellec, Lax, and Phillips 2010). While it is straightforward to measure levels of government spending across different policy areas, we address a key limiting factor for this literature: the absence of reliable measures of public preferences for levels of spending across policy domains.

Second, public opinion on economic issues in many countries is often treated as though it varies primarily according to a single ideological dimension that captures the classic left–right divide over the size of government (Caughey, O'Grady, and Warshaw 2019; Meltzer and Richard 1981). We show that voters do not simply favor more or less government spending (or taxation) overall, but rather favor higher spending in some areas and lower spending in other areas. This claim is distinct from the fact that politics in many countries features a multidimensional mix of economic and social issues: Even when focusing solely on fiscal policy, improved measurement reveals that voters' preferences are not closely approximated by a unidimensional structure. These results speak to recent work on coalitions of support for the reorientation of welfare spending from traditional categories to

---

[1] A separate criticism of using stated preferences to elicit preferences is that surveys ask inherently hypothetical questions of respondents. Our approach does nothing to improve on this limitation, but given the ubiquity of survey-based approaches and the different problems associated with behavioral measurements of budget preferences in public opinion, improvements within the survey method are valuable.

[2] Government spending can also be financed through increased borrowing, but here we focus respondents' attention on the key trade-offs between tax and spending. Our approach easily permits changing deficit levels.

new ones (Häusermann 2010). However, our UK respondents do not divide in a budgetary competition between social investment (education spending) and consumption (pensions; Busemeyer et al. 2018). Instead, our respondents prefer to increase both at the expense of higher taxes and spending reductions elsewhere.

Finally, scholars in comparative political economy are increasingly attentive to budget trade-offs in social and fiscal policymaking at times of austerity. An important body of work asks about voter preferences on the trade-off between spending cuts and tax increases, if borrowing is off the table so that the government budget constraint is hard (Bremer and Bürgisser 2019; Hübscher, Sattler, and Wagner 2020). Our approach provides an additional point of evidence that tax increases may be supported over spending cuts.

What Do We Know about Budget Attitudes and How to Measure Them?

Public preferences over government spending have long been an object of research due to their centrality to a number of important substantive questions. How voters want the government to allocate funds is critical in understanding the functioning of democratic systems—for example, in the literature on policy responsiveness, and in questions on the political economy of government spending more broadly. All of these substantive inquiries, when using data on public opinion on spending, make choices about the measurement of these preferences. These choices (implicitly) propose some solution to the trade-off between simplicity for the gauging of opinion, along with realism in the budget preferences that are revealed. Leaving these measurement questions unexamined has led, in part, to inconsistencies in the kinds of preferences attributed to voters, who have then in turn been characterized as inconsistent. Examining these debates highlights the need for our contribution here, echoing other recent attempts to generate explicit and absolute public positions on government spending allocations to different areas of the budget (Bonica 2015; Branham and Jessee 2017).

First, understanding public opinion on government spending has been a central concern of scholars interested in the degree of government intervention in the economy, in particular in its actions to offset inequality. In this tradition, the centrally important preferences are those over the level of government spending overall, or aggregated across a large set of policies (Kelly and Enns 2010; Svallfors 2012); spending on particular social policies with important distributive implications, such as unemployment insurance (Rehm 2011); and on the broader question of support for redistribution (Cavaillé and Trump 2015; McCall and Kenworthy 2009; Rueda and Stegmueller 2019). The core themes of this work center on the role of material self-interest in shaping these preferences, as well as overall levels of support for egalitarian spending. In the latter case, results indicating generally high levels of support for public spending and redistribution raise questions as to why policy has not followed along these lines, with some pointing to the failure of these questions to capture the tax costs of such budget changes.[3]

The second area of attention compares budget allocations to public opinion as an indicator of policy responsiveness to preferences, and of opinion responsiveness to spending (Soroka and Wlezien 2005, 2009; Wlezien 1995). These studies typically measure spending preferences in relative terms; for example, "Are we spending too much, too little, or about the right amount on the military, armaments, and defense?" These questions are simple for respondents to evaluate, but they do not translate directly into specific preference positions over a full budget. Importantly, this kind of measure makes no explicit allowance for trade-offs in the budget process (a cost of higher spending in one area is reductions in another, or increased taxation). Nor does this allow for any prescription as to the size of any preferred change—a large majority of people in favor of an increase is not equivalent to a majority in favor of a large increase.

These approaches have led to a body of public opinion data that reflects more nuance than simple pro-spending versus anti-spending views (Page and Jacobs 2009), but which equally may reflect variation in the salience of the trade-offs that must be made in budget making in the real world, and the perceived feasibility of viable changes when voters may be misinformed about the status quo.[4] When taken together, the cumulation of these simplified questions across different areas of the government budget—including implications for overall rates of taxation—tends not to provide a clear and coherent view of budget preferences.

---

[3]Much of this work also considers public preferences over the revenue-raising side of the budget, as well as the spending side, but usually taken separately rather than linked via a budget constraint (Bartels 2005; Gilens 2012). We consider only the overall tax burden implications of different allocations on the spending side, but an analogous exercise in examining how the public would prefer government revenues to be raised is a natural application of the same underlying measurement strategy that we introduce here.

[4]For example, Page and Jacobs (2009) discuss voter "confusion" about the extent of foreign aid spending, and thus the implications for the impact of cuts to aid on the overall budget.

Indeed, the results of these strategies of measurement have been divergent enough to suggest that voters want something for nothing, and are not equipped to deal sensibly with trade-offs, although the evidence for such inconsistent preferences "is mixed" (Citrin 1979, 128). Voters may have beliefs that provide good reason for denying a tight trade-off exists—in particular, the perception of waste (Citrin 1979; Williamson 2017). If there is waste to be trimmed, a desire to increase spending without increasing taxes is not inconsistent. In other cases, voters may not accurately perceive the costs and benefits on each side of the trade-off. Scholars have explained a lack of willingness to pay taxes by the inadequate visibility of the spending they support (Downs 1960; Mettler 2011), and support for spending programs by the inadequate visibility of the taxes required to support them (Winter and Mouritzen 2001). Alternatively, patterns of public opinion that appear inconsistent at the aggregate level can arise from preferences that are rational at the individual level, where taxpayers prefer higher levels of spending financed by increases in the taxes paid by other people (Edlund and Sevä 2013).

When survey prompts make budget trade-offs more explicit, respondents provide more consistent profiles of responses. The most complete existing study of budget preferences under trade-offs was fielded in the pilot to the 1996 American National Election Study, presenting respondents with choices on exhaustive and mutually exclusive categories (taxation, deficits, and domestic and defense spending) that incorporated the relevant trade-off and asking about the level of support for policy packages that combined increases in one area with offsetting decreases elsewhere. Hansen (1998) finds that these data reveal a high degree of consistency in budget preferences.

Thus, it is easy to demonstrate that people do not give responses that obey budget constraints when given questions that neither enforce nor encourage obeying such constraints. It is not inconsistent or irrational to prefer both high spending and low taxes in a situation where there is no reason you cannot want both. Most people want both, all else equal. The point of the budget constraint is that all else cannot be equal. But if you are answering a survey that only mentions one at a time, or if you are left free to implicitly externalize tax costs to other people, the budget constraints do not obviously bind. This is not a failure of the respondents, but a failure of the survey instrument. If our goal is to elicit public preferences over feasible budgets, it makes sense to ask respondents about feasible trade-offs.

State-of-the-art measurement in this area has thus recently moved in the direction of richer survey prompts and spatial modeling, in order to better elicit information about trade-offs. First, the logic leads very naturally to the idea of simply asking each respondent to set spending levels across all budget categories. This approach is taken by Bonica (2015), who uses an online budgeting task that allows respondents to reallocate the U.S. federal budget from its existing baseline (holding tax and deficit constant) across 21 spending areas. This approach has been adopted in a number of recent studies (D'Attoma, Tuxhorn, and Steinmo 2018; Hübscher, Sattler, and Truchlewski 2020). However, solving the full budget allocation represents a highly challenging task for respondents, motivating the search for a more accessible survey instrument. Second, Branham and Jessee (2017) use U.S. General Social Survey data on "too much," "about right," and "too little" assessments of spending across 18 spending areas, and they use a spatial model to map responses onto a distribution of public preferences in a unidimensional space. This makes maximal use of this kind of (relatively accessible) survey question, but it does not provide quantitative estimates of the size of preferred changes and (as the authors note) does not make respondents face trade-offs directly (though the spatial dimension implicitly describes a range of different preferred trade-offs).

We build on Bonica (2015) by targeting an absolute and quantitative preference for spending change, with full attention to the multidimensional trade-offs involved. But our concerns about the difficulty for respondents in constructing a full budget suggest our intermediate solution in which we ask about narrower trade-offs involving a few spending categories at a time. By exploring the space of possible trade-offs, we can use a spatial modeling approach like that of Branham and Jessee (2017) to reassemble responses into a budget.

Finally, our approach builds on recent empirical developments in the analysis of trade-offs in general. These have come increasingly under investigation with the use of conjoint experiments. Conjoint choice analyses have been used to ascertain preferences over various multidimensional policy areas, including tax progressivity (Ballard-Rosa, Martin, and Scheve 2017), climate change agreements (Bechtel, Genovese, and Scheve 2017), and immigration (Hainmueller and Hopkins 2015), to name but a few.

There have already been attempts to use forced-choice designs to investigate budget preferences. In Bremer and Bürgisser (2019) and Hübscher, Sattler, and Wagner (2020), respondents are asked to choose a package of fiscal proposals that differ in terms of spending, taxation, and government debt. In each of these approaches, the direction of the changes are specified—increases or decreases in taxes and spending of different types, which may be "small" or "large." Thus,

preferences over tax levels and spending within this directionally balanced budget are elicited. These studies show little (if any) differentiation across spending types in terms of the effect of cuts on support for the package (e.g., cutting healthcare is equally as unpopular as cutting pensions), but this is a hard task for this empirical design as there is no obvious metric by which to convert "small" cuts in one spending area to "small" cuts in another. Our approach makes these trade-offs more concrete by requiring not just directionally but also arithmetically balanced budgets.

Our approach draws on the intuition that a multivariate, forced-choice experiment like a conjoint experiment is useful for understanding preferences over multidimensional objects like budgets. But the standard conjoint design with independent treatments is ill-suited to the quantitative trade-offs involved in budgets; budget trade-offs are, by their nature, not independent. Increasing one spending area requires changing other spending areas or taxation to maintain balance. These changes are perfectly multicollinear: Knowing the changes made in $K-1$ categories of tax and spending determines the remaining category exactly. This makes the resulting data not only incompatible with the standard linear regression analysis of conjoint experiments, but also with the theoretical logic of calculating average marginal component effects (Hainmueller, Hopkins, and Yamamoto 2014). In order to explore these budget trade-offs, we need a different sort of multivariate forced-choice experiment, along with a new strategy to analyze the resulting data.

Experimental Design and Data Collection

We fielded an experiment in October 2018 where we presented respondents with information mirroring the taxpayer summaries distributed by the UK government. These summaries, which are distributed to every UK taxpayer at the end of the financial year, indicate how individuals' income tax and national insurance contributions are allocated across a variety of public spending categories. The summaries therefore provide itemized accounts of government spending over the previous year, and they are personalized for each recipient by scaling them to that individual's total tax level. Further detail on these summaries is provided in the supporting information (p. 1).

We used the tax summary spending allocation from fiscal year 2016/17 as a baseline and then—for each respondent—randomly altered the amount of spending in three spending categories and (in some cases) the

overall tax level. Respondents were first presented with an introduction screen, which outlined the basic idea of the task, and provided information on how government spending is currently allocated across the 15 spending categories. We expressed all information about tax and spending levels with reference to the income level of an individual who earned £23,200 each year, approximately the median income of a UK taxpayer in 2016.

Respondents were then asked to consider a series of comparisons between randomly generated proposed tax and spend levels and the status quo allocation. All proposed changes involved maintaining deficit neutrality; proposed changes in spending and tax always exactly offset. In one-third of the proposals, there was no change in tax, only shifts between spending categories. The prompt used in the experiment is given in Figure 1. We provided respondents with current and proposed spending levels on each selected category in pounds, and also provided information on the implied change in spending and tax in both pounds and percentages. In addition, we provided text to clarify which spending areas would see increased or decreased spending under the proposed plan relative to the status quo, and which areas would see no change in spending. Respondents were then asked to indicate whether they preferred the current levels of spending or the proposed changes, or whether they were not sure. Each respondent faced five different choice tasks, each time comparing a randomly selected spending proposal to the status quo. Details of how we generated the distribution of proposed changes are provided in the supporting information (p. 3).

We suggested above that part of the logic of this approach is accessibility to respondents compared to a full allocation task. The simplest test of respondent engagement with our prompt is to assess how many people give "I am not sure" responses, and in particular how many give this response to all five proposed sets of changes that they saw. Of 3,533 respondents overall, 52% gave zero "not sure" responses, 16% gave one, 9% gave two, 4% gave three, 4% gave four, and 16% gave five. We would expect that even fully engaged respondents would give this response to some fraction of the prompts, but the relatively large number of respondents giving this response to five items (versus three or four) confirms that some respondents did not engage with the questions. Respondents who gave five "not sure" responses spent, on average, 56% less time on the survey than respondents who provided zero "not sure" responses. Those with lower levels of education were more likely to not express preferences on any comparison: 11% among those with university degrees versus 19% among those without. These are moderate levels of differential nonresponse that are typical of those seen in political survey data (Berinsky 2008).

FIGURE 1 Experiment Prompt

Note: The bottom five rows group respondents by their 2017 general election and 2016 Brexit referendum vote combinations.

Overall, this experiment gives us rich information about how citizens make trade-offs between different spending areas and tax, but it is not straightforward to analyze. It is not a traditional conjoint experiment because the experimental treatments are not merely nonindependent, but perfectly multicollinear. We cannot fit regressions for the responses as a function of the treatments because of this multicollinearity: The "all else equal" logic of regression modeling is ill-suited to data where all else is never equal by design. Instead, we use a structural model that assumes respondents have spatial preferences over spending and tax levels and make choices accordingly. Given these assumptions and the observed responses, we can infer the underlying distributions of preferences most likely to have generated the data.

A Model for Respondent Choice

To build a model for respondents' choices, we need to make some assumptions regarding how they will trade off changes in spending and taxation. We adopt an ad-

ditive quadratic loss model for deviations from respondents' preferred spending level in each spending area, which is consistent with typical approaches in the spatial preference modeling literature (Clinton, Jackman, and Rivers 2004) and which proves mathematically convenient for defining our estimator below. Respondents i make a choice between the proposed alternative (A) and the status quo (S), with an option to say they are not sure (NS) which they prefer. We assume that they do so as a function of the latent utility of the two alternatives, with "not sure" corresponding to cases where their utility difference is small.[5]

$$Y_i = \text{proposed alternative if } u_{iA} - u_{iS} > \gamma_A,$$

$$Y_i = \text{not sure if } \gamma_S \leq u_{iA} - u_{iS} < \gamma_A$$

and

$$Y_i = \text{status quo if } u_{iA} - u_{iS} \leq \gamma_S.$$

[5]In order to identify the latent utility scale relative to the response, we assume that the thresholds that respondents apply for selecting either the proposed alternative or the status quo, as opposed to "not sure," are symmetric around zero and set to $\gamma_A = +1$ and $\gamma_S = -1$.

We assume that respondents i have quadratic loss in deviations of the status quo $S = (S_1, \ldots, S_M)$ and the proposed alternative $A_i = (A_{i1}, \ldots, A_{iM})$ from their preferred point $\Psi = (\psi_{i1}, \ldots, \psi_{iM})$, and that this is weighted per dimension $j \in 1, \ldots, M$ by a factor λ_j. Further discussion of these choice model assumptions is provided in the supporting information (p. 7).

$$u_{iA} = -\sum_j \lambda_j \left(A_{ij} - \psi_{ij} \right)^2,$$

$$u_S = -\sum_j \lambda_j \left(S_j - \psi_{ij} \right)^2$$

and

$$u_{iA} - u_{iS} = \sum_j \lambda_j \left(S_j^2 - A_{ij}^2 \right) + 2\lambda_j \left(A_{ij} - S_j \right) \psi_{ij}.$$

The λ_j weights mean that the model allows for respondents putting similar weights on £ deviations in all categories, or similar weights on percentage changes from baseline spending, or some other weighting of categories. The values of S_j and A_{ij} are data, and we have plenty of data with which to estimate λ_j for each of the dimensions. The thing we neither know nor can estimate with useful precision is a single respondent i's preferred allocation in each spending area j, his or her ideal point ψ_{ij}.

Estimating the Distribution of Preferences

Although we do not have enough data to measure ψ_{ij} at the individual level, we can estimate the distribution of these individual-level ideal points in the population. We assume that ψ_i follows a multivariate normal distribution:

$$\psi_i \sim N(\mu, \Sigma),$$

where μ is a vector of length M giving the average preferred allocation in each spending area, and Σ is the covariance matrix of individual respondents' preferred allocations around that average with elements $\Sigma_{jj'} = \rho_{jj'} \sigma_j \sigma_{j'}$. The utility difference $u_{iA} - u_{iS}$ then defines an affine transformation on this multivariate normal distribution, and it thus implies a univariate normal distribution for the utility difference between the two alternatives faced by a given respondent:

$$u_{iA} - u_{iS} \sim N\left(v_i, \omega_i^2 \right),$$

$$v_i = \sum_j \lambda_j \left(S_j^2 - A_{ij}^2 \right) + 2\lambda_j \left(A_{ij} - S_j \right) \mu_j$$

and

$$\omega_i^2 = \sum_j \sum_{j'} \rho_{jj'} \sigma_j \sigma_{j'} \left(2\lambda_j \left(A_{ij} - S_j \right) \right) \left(2\lambda_{j'} \left(A_{ij'} - S_{j'} \right) \right).$$

Given our definition of the response at the outset, this yields the following response distribution, conditional on parameters, for the three responses:

$$p(\text{status quo}) = \Phi\left(\frac{\gamma_S - v_i}{\omega_i} \right),$$

$$p(\text{not sure}) = \Phi\left(\frac{\gamma_A - v_i}{\omega_i} \right) - \Phi\left(\frac{\gamma_S - v_i}{\omega_i} \right)$$

and

$$p(\text{proposed alternative}) = 1 - \Phi\left(\frac{\gamma_A - v_i}{\omega_i} \right).$$

This model has a moderate number of free parameters—13 μ_j, 13 λ_j, 13 σ_j, and 78 $\rho_j = 117$ total—but we have 17,665 responses with which to estimate those parameters (about 150 per parameter). Because our experiment can most meaningfully test spending levels that are included in the distribution of provided treatments, and that distribution reflected our ex ante expectations, we put a prior on our estimated average preferred spending levels that corresponds to that treatment distribution. This prior is $\mu_j \sim N(S_j, sd_j)$, where S_j is the status quo/baseline spending in each area as described above, and sd_j is the standard deviation of the nonzero treatments provided on each spending area. Replacing this informative prior with a flat prior has little effect on the estimates of μ_j.

It is straightforward to extend the model above to enable the description of variation in individual preferences as a function of covariates. Instead of estimating a single vector of average preferred spending levels μ, we instead model individual preferences by modeling the mean vector of the multivariate normal distribution as

$$\psi_i \sim N(\alpha + \beta X_i, \Sigma)$$

for some $1 \times K$ vector of covariates X_i describing each respondent i and estimate a $M \times 1$ vector of intercepts α and a $M \times K$ matrix of coefficients β that describe variation in preferences as a function of each of the K covariates on each of the M spending dimensions. For these models, we impose the same prior on α that we put on μ in the model without covariates, along with a common shrinkage prior on all β parameters proportional to the category size $\beta \sim N(0, S_j \sigma_\beta)$ to avoid overfitting small subgroups.

We estimate the models with and without covariates using Hamiltonian Monte Carlo as implemented in Stan (Carpenter et al. 2016). Our presented results consist of mean posterior point estimates and central 95% intervals

FIGURE 2   Average Preferred Change in Spending (% of Current Level) with 95% Posterior Intervals
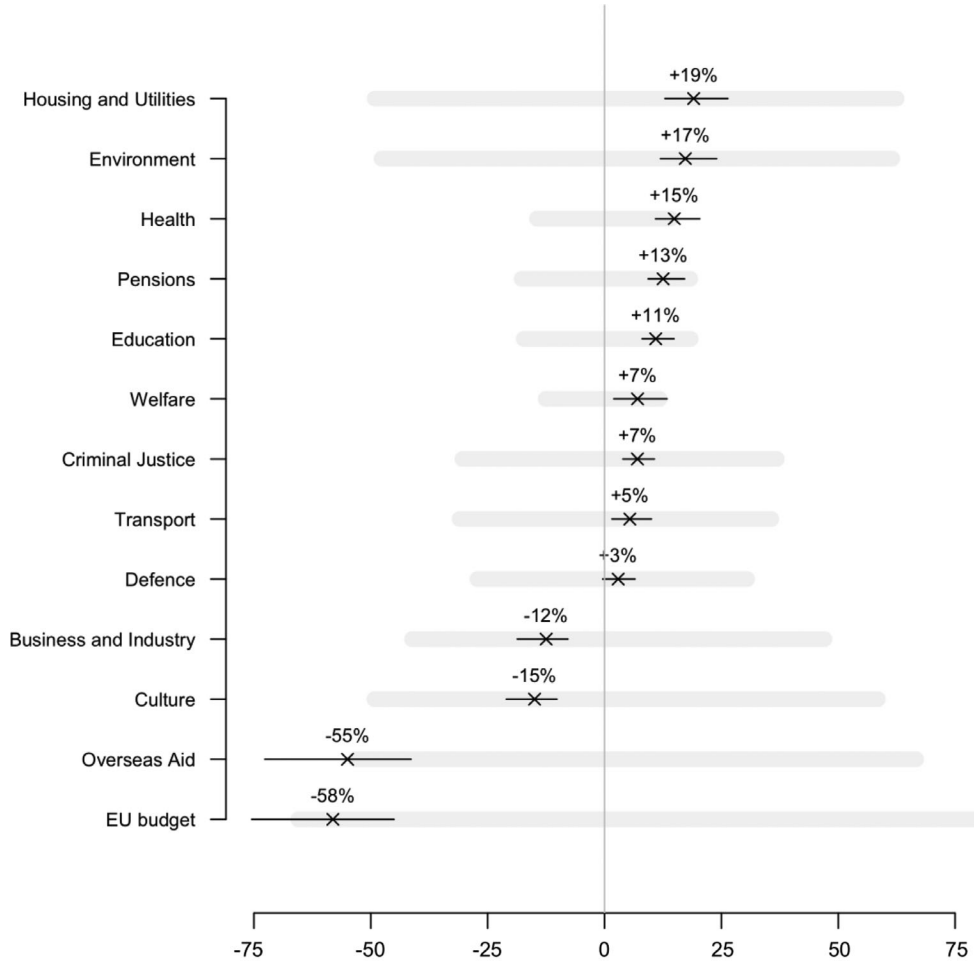
Note: The light gray bands show the central 95% range of the treatment distribution for that spending area.

based on five parallel simulation chains of 10,00 iterations (after a 250-iteration warm-up) for each model. In order to maximize the degree of representativeness with respect to the UK population, we use demographic survey weights provided by YouGov via a quasi-likelihood approach. The estimates are very similar without using the weights.

Results

Average Preferences

Figure 2 presents the average preferred change in spending for each of the spending categories in our experiment, expressed in percentage terms (where current levels are the baseline). UK respondents, on average, endorse spending increases of greater than 10% in educa-

tion, pensions, health, the environment, and housing; we also see smaller increases in transport, welfare, and criminal justice. By contrast, UK respondents favor budget cuts in relatively few spending areas: business and industry, culture, overseas aid, and UK contributions to the European Union (EU) budget. People want more spending in the four categories with the largest baseline levels: education, pensions, health, and welfare. Likewise, the largest preferred spending cuts—to overseas aid and EU budget contributions—are in the two smallest categories we included in the experiment. One consequence of this is that the average preferred overall spending levels are 1.08 times the current spending levels in the categories where we proposed changes. Once we take into account the two budget categories in which we did not test changes—payments on the national debt and government administration—this implies that respondents

would be willing to accept an overall tax increase of 7% to pay for these increases in spending.[6]

Model Checks

In the supporting information (p. 8), we look in detail at the values of the auxiliary parameters λ, σ, and ρ, as well as examining the several alternative specifications of the model. We briefly summarize some key findings here. The dimensional weights λ_j are very close to inversely proportional to the baseline spending levels; the correlation coefficient of the λ_j parameters with $1/S_j$ is $r = 0.985$. This indicates that respondents tend to penalize alternatives in terms of percentages, not pounds (recall that we present both in the experiment). The ρ_j indicates that preferences across spending areas are correlated in expected ways, given how UK politics is organized, but not very strongly. The most positively correlated preferences are those for the core social welfare categories of welfare, education, and health.

In the supporting information (p. 9), we also report the estimates of μ from our main model compared to (1) the same model with a flat prior on the μ_j, (2) a model where we assume quadratic loss in log spending rather than spending, (3) a model where we drop respondents who gave "I am not sure" responses to all five proposed changes, (4) a model where we estimate responses separately for each of those five rounds of responses, and (5) a model where we use only responses to proposals that involve no tax change. Overall, the estimates are robust to these model and data variations, telling a consistent story about respondents' relative preferences for spending in different areas.

Cross-Sectional Preference Variation

As we described earlier, the average preferences of the public can be further disaggregated in terms of covariates. We present a series of simple models (Figure 3) with single categorical variables for age (18–34, 35–54, 55+), university degree (no, yes), household income (below £30,000, above £30,000, refused), and combinations

of votes in the 2016 EU referendum during the 2017 UK general election (Conservative-Leave, Conservative-Remain, Labour-Leave, Labour-Remain, other). In addition, we fit a multivariate model with age, degree, income, 2016 vote, and 2017 vote in order to check whether the univariate patterns that we find for some of these variables might reflect the correlations between the covariates. We present the results from the multivariate model in Figure 4.

Figure 3 (simple models) and Figure 4 (the multivariate model) provide results that both confirm the face validity of the estimates and reveal some interesting patterns that would not be obvious given commonly held stereotypes about contemporary British politics. Regarding face validity, we see the basic strong political relationships that we would expect. The 2017 Conservative voters want less spending and lower tax overall than 2017 Labour voters. Labour voters want more spending than Conservative voters on most categories of spending, especially welfare, education, health, the environment, and housing. Conservative voters want more defense spending than Labour voters. Leave voters want far lower spending on the EU (and overseas aid) than Remain voters, but also less on the environment, welfare, transport, culture, and housing. None of these relationships are surprising—for many, we would have reason to worry about the validity of the method if it had found otherwise.

What might be more surprising to observers of British politics are the associations of respondents' tax and spending preferences with age. Age-related differences in party choice are not new features of UK politics, but the 2017 UK general election saw a historically remarkable age gradient in voting, with Labour winning those aged 18–34 by a 60% to 27% margin and the Conservatives winning those aged 55 and older by 54% to 31%. This gap emerged in the context of an election where the Labour Party campaigned on a manifesto that pledged to significantly increase spending in several areas of the government budget and to increase both income and corporation taxes. Several commentators took the increase in support for Labour among the young to emphasize the new importance of age cleavages in UK politics,[7] with young supporters of the Labour Party described recently as "millenial socialists."[8]

---

[6]While the estimated average preferred changes are within the range of treatments we tested for every spending category, their cumulative effect on the tax level is not. This occurs because our data tell us that respondents prefer spending increases in many areas when they are considered three at a time, but not whether they would be willing to increase taxes to fund all of those increases at once. The quadratic loss preferences we specify imply that they would be willing to pay for several increases, and our model estimates that the size of those increases would require a tax hike that is outside of the tax changes that we proposed in the experiment.

[7]See "The Biggest Political Divide in Britain Is Age," *Economist*, June 8, 2017, https://www.economist.com/speakers-corner/2017/06/08/the-biggest-political-divide-in-britain-is-age.

[8]See "Millennial Socialism," *Economist*, February 14, 2019, https://www.economist.com/leaders/2019/02/14/millennial-socialism.

FIGURE 3 Average Preferred Changes in Each Spending Area and Overall Tax (Columns) for Various Subgroups (Rows)

Defence	Welfare	Education	Health	EU budget	Business and Industry	Pensions	Criminal Justice	Environment	Transport	Culture	Housing and Utilities	Overseas Aid	Tax	
0	+11	+16	+19	-59	-16	+13	+7	+23	+7	-12	+21	-51	+11	University Degree
+5	+5	+9	+12	-58	-11	+13	+7	+14	+5	-17	+18	-56	+7	No University Degree
+3	+9	+12	+15	-56	-10	+12	+7	+17	+3	-12	+18	-52	+9	Female
+3	+5	+10	+16	-59	-14	+13	+7	+18	+8	-18	+21	-57	+8	Male
+1	+7	+12	+14	-58	-12	+12	+6	+19	+6	-14	+20	-54	+8	Income Refused
+4	+6	+12	+17	-58	-14	+15	+7	+17	+6	-16	+18	-56	+9	Income > £30k
+4	+8	+10	+10	-58	-11	+13	+8	+17	+4	-15	+20	-55	+7	Income < £30k
+14	+15	+14	+22	-67	-7	+22	+13	+18	+8	-24	+29	-63	+14	55+
-2	+9	+11	+15	-56	-15	+10	+8	+18	+4	-13	+12	-48	+8	35-54
-9	-3	+8	+10	-42	-19	+4	-1	+16	+4	-4	+14	-43	+2	18-34
-1	+3	+10	+14	-43	-18	+10	+6	+14	+5	-13	+6	-36	+6	Other Vote Combinations
+9	+16	+8	+17	-51	-14	+16	+10	+16	+15	-24	+17	-53	+12	2017 Con 2016 Remain
-11	+27	+18	+23	-19	-14	+15	+4	+37	+13	-2	+42	-14	+18	2017 Lab 2016 Remain
+14	-12	+9	+15	-90	-10	+17	+12	+2	-1	-24	+17	-81	+3	2017 Con 2016 Leave
+15	+13	+12	+18	-59	-2	+17	+13	+22	+3	-20	+26	-68	+12	2017 Lab 2016 Leave

Note: The bottom five rows group respondents by their 2017 general election and 2016 Brexit referendum vote combinations.

However, in our survey, younger respondents prefer substantially *lower* levels of taxation and spending overall. This is true both holding constant the 2017 vote (Figure 4) and also unconditionally (Figure 3). In the simple group comparison (Figure 3), the average increase in overall spending and tax preferred by those over 55 is 14%, versus just 2% for those under 35 years old. Young people endorse lower spending levels than older people in every spending category except the contribution to the EU budget, culture, and overseas aid, which are all among the smallest budget areas. This pattern is in stark contrast to the widely held understanding in UK politics that younger cohorts are more amenable to greater public expenditure and taxation than older voters,[9] though it is consistent with recent cross-national evidence on generational fiscal preferences (O'Grady 2020). Unless there is some very large age-associated bias in our measurement strategy, which we think is unlikely, it is the older voters
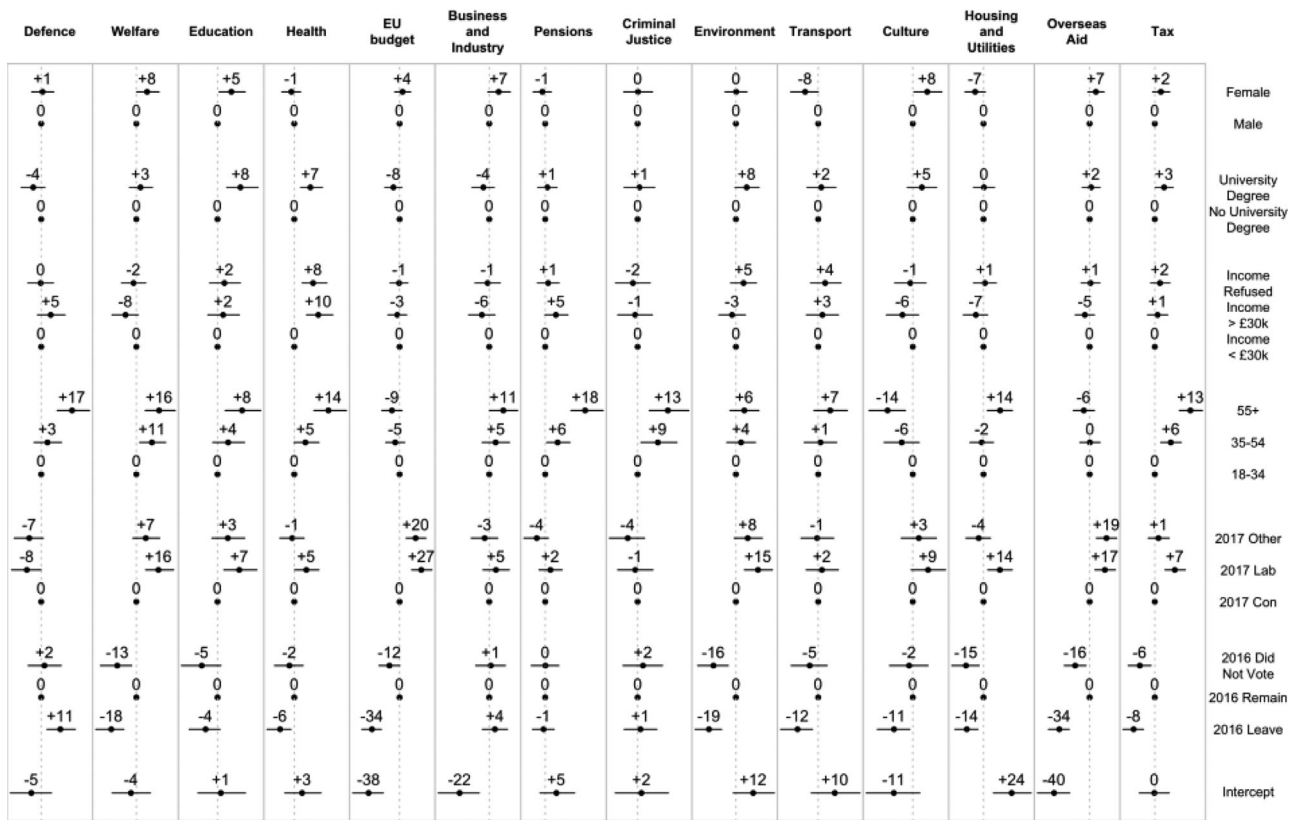
who want to substantially increase government spending, not the younger voters.

Overall, a key finding of this analysis is that citizens' spending preferences are multidimensional rather than tightly structured by a single preference dimension. In the supporting information (p. 11), we report details of principal components analysis on the estimated distribution of preferences from the model. The first dimension is oriented from Conservative-Leave to Labour-Remain, associated with wanting more spending on everything except Defense, Business, Pensions, and Criminal Justice. The second dimension is oriented from Conservative-Remain to Labour-Leave, associated with wanting more spending on everything except the EU budget, culture, and overseas aid.

We do not see a strong first dimension, followed by a rapid decline in preference variation explained. The Conservative-Leave to Labour-Remain dimension is more predictive of spending preferences across areas than any other dimension, but not much more predictive. Thus, we find that preferences over spending in different categories vary in a complex way across individuals

[9]Grasso et al. (2019) has some suggestion of this finding, but the measures used in that research are less precisely focused on spending.

FIGURE 4 Multivariate Model Coefficient Estimates as a Function of Covariates

	Defence	Welfare	Education	Health	EU budget	Business and Industry	Pensions	Criminal Justice	Environment	Transport	Culture	Housing and Utilities	Overseas Aid	Tax
Female	+1	+8	+5	-1	+4	+7	-1	0	0	-8	+8	-7	+7	+2
Male	0	0	0	0	0	0	0	0	0	0	0	0	0	0
University Degree	-4	+3	+8	+7	-8	-4	+1	+1	+8	+2	+5	0	+2	+3
No University Degree	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Income Refused	0	-2	+2	+8	-1	-1	+1	-2	+5	+4	-1	+1	+1	+2
Income > £30k	+5	-8	+2	+10	-3	-6	+5	-1	-3	+3	-6	-7	-5	+1
Income < £30k	0	0	0	0	0	0	0	0	0	0	0	0	0	0
55+	+17	+16	+8	+14	-9	+11	+18	+13	+6	+7	-14	+14	-6	+13
35-54	+3	+11	+4	+5	-5	+5	+6	+9	+4	+1	-6	-2	0	+6
18-34	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2017 Other	-7	+7	+3	-1		+20	-3	-4	-4	+8	-1	+3	+19	+1
2017 Lab	-8	+16	+7	+5		+27	+5	+2	-1	+15	+9	+14	+17	+7
2017 Con	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2016 Did Not Vote	+2	-13	-5	-2	-12	+1		+2	-16	-5	-2	-15	-16	-6
2016 Remain	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2016 Leave	+11	-18	-4	-6	-34	+4	-1	+1	-19	-12	-11	-14	-34	-8
Intercept	-5	-4	+1	+3	-38	-22	+5	+2	+12	+10	-11	+24	-40	0

Note: Estimates are coefficients for average preferred spending in percent of the current baseline for each spending area and overall tax (columns). Rows indicate covariates.

and demographic groups that does not reduce to some groups simply being in favor of more or less spending, or more left-wing versus right-wing spending priorities. That preferences are not simply unidimensional should not be surprising given what we know about the high level of individual-level idiosyncrasy in citizens' political preferences in other domains (Broockman 2016; Lauderdale, Hanretty, and Vivyan 2018), even though the idea of a dominant left–right economic dimension is widespread in the political science literature.

Validation

Ideally, a model that infers a distribution of preferences of individual voters based on their survey responses would be able to predict survey responses to similar budget trade-off questions out-of-sample. We evaluate the out-of-sample predictive performance of our modeling approach in two ways. First, we conduct an *internal* validation, where we evaluate the ability of the model estimates to predict responses to the experimental prompt that are held out from estimation of the model. Second, we conduct an *external* validation, where we evaluate the ability of the model estimates to predict responses to a novel budgetary proposal in a new survey.

Internal Validation

Our internal validation is a fivefold cross-validation. We fit our baseline model and our full covariate model five times, each time holding out one round of responses to the experiment. We then construct predicted probabilities for responses to the held-out round from the model fit that excluded that round, generating predictions for all responses to the experiment, none of which were fit on the predicted responses. Both the baseline and covariate models predict variation in the probabilities of each response across proposals because some changes (e.g., reducing health spending) tend to decrease support for proposals, whereas others tend to increase it across

FIGURE 5 Cross-Validation Predictive Performance by Response Level
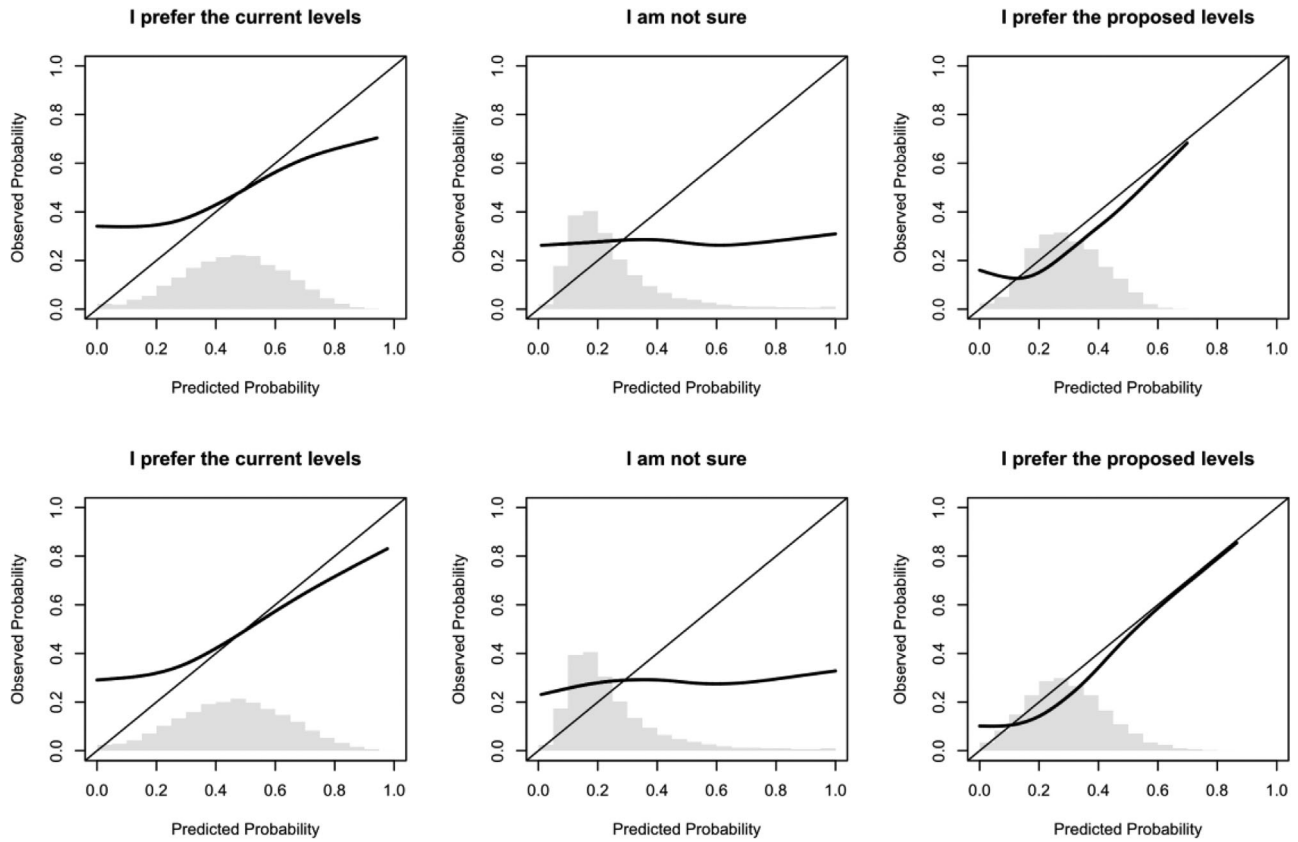
Note: The analysis uses the model without (top row) and with (bottom row) respondent-level predictors. Each plot shows a spline fit for the observed response on the prediction overlaid on the distribution of predicted probabilities.

all respondents. The covariate model is additionally able to predict variation across respondents, holding the proposal fixed.

Figure 5 shows that while the model predicts the probabilities of respondents endorsing the proposal reasonably well, with or without respondent-level predictors, it does less well at predicting endorsement of the status quo spending levels and does quite poorly at predicting the "I am not sure" response. By implication, the model is not doing well at distinguishing between endorsing the status quo and the "not sure" response. This is not a problem related to overfitting: These plots are indistinguishable from when we examine in-sample fit. The functional form of our model cannot fully describe the patterns of responses that we see as a function of the proposals.

The primary reason for the lack of predictive fit is that our model does not capture all "not sure" mechanisms well. This has a limited consequence for the estimates of the model; when we construct estimates excluding respondents who gave "I am not sure" responses to all five proposals, there is negligible change in the es-

timates (see p. 12 of the supporting information), even though this drops about half of the "not sure" responses in the data set. The fact that there are a substantial number of such respondents, however, highlights that our model can only make sense of nonresponse as the result of narrow indifference between the status quo and the proposal. In reality, "I am not sure" responses are likely to arise from several mechanisms, including this sort of narrow indifference, but also including strong cross-pressure (e.g., a strongly liked change combined with a strongly disliked change) that respondents find difficult to weigh, or general disengagement (e.g., respondents who always say they are not sure). Future research could leverage this experimental design to examine the relative extent to which, and among which groups of respondents, "not sure" responses arise from these different mechanisms.

External Validation

Because we estimate not only the average preferences of UK citizens (via the μ parameters), but also their

distribution and relative weight (via the σ, ρ, and λ parameters), we can estimate what proportion of respondents we would expect to support a proposal to change all spending and tax levels to the average preferred level, versus the status quo. According to our structural model, we calculate that 72.7% of respondents should prefer the proposal to the status quo (SE: 0.028; 95% interval: 0.681–0.792) and 27.3% should prefer the status quo, according to whether they have positive or negative latent utility differences $u_{iA} - u_{iS}$. In the original experiment, status quos were preferred nearly twice as often as the proposed changes. Have we in fact found a set of changes that is substantially preferred to the status quo, even when most proposed changes were not preferred in the original experiment?

Our external validation study was also conducted by YouGov, using the same sampling methods from their UK panel, and we asked a single question of 3,000 respondents. We presented respondents with the status quo figures for each spending category and the overall tax rate from the baseline tax summary, and then we asked them to compare that with a proposal that involved changing all spending categories to the average preferences estimated from our model. The entire prompt, as delivered, is provided in the supporting information (p. 6). We preregistered this validation study and our standards for a successful validation before data collection.[10]

The results of the validation were as follows. Of the population-weighted respondents, 39% indicated that they supported the proposed changes, versus 25% who indicated that they preferred the current spending levels; 36% indicated that they were not sure. Using the estimator that we preregistered to maximize comparability, we calculate that this response distribution corresponds to 58.3% of respondents preferring the proposed changes to the status quo versus our structural model prediction of 72.7%.

According to our preregistered assessment criteria, this counts as a "moderate success." We were able to use the experiment to identify a profile of changes that was supported by substantially more respondents than opposed it, even though the level of support was not as high as the structural model implied. In the preregistration, we identified several reasons to expect that fewer respondents would endorse this proposal than our structural model implied, including status quo bias that did not reflect the actual content of proposals, the larger tax hike involved in this proposal, and the greater difficulty of the task of contemplating changes to 13 spending cat-

egories rather than three (Barnes, Blumenau, and Lauderdale 2019).

Moving beyond the basic evaluation set out in the preregistration, we poststratify the structural model estimates onto the validation survey respondents and then characterize the relationship between predicted support levels and actual support for the proposal among those respondents. Figure 6 shows the results of this analysis. Although the observed levels of support for the proposal are consistently lower than the predicted levels, they are strongly associated with those predictions, with a slope close to one and only minor nonlinearity in the relationship (left panel). When the individual respondents are aggregated into the same categorical groupings used previously (right panel), we see that the demographic groups that the structural model indicated should be more favorable toward the proposal are in fact more favorable toward that proposal.

Validation Summary

Do these internal and external validation results undermine the estimates we produced of the average preferred spending (and thus tax) levels? To the extent that these results indicate that respondents do not exactly approximate the choice logic underlying our structural model, this does imply lesser confidence in the overall spending *level* of the estimates as the true "ideal" for the UK population. Nonetheless, we are able to make valid out-of-sample predictions for which three-category change proposals would be more or less favored by respondents. We are also able to identify a full alternative budget that received substantially more support than opposition, despite high survey item complexity, potential for loss aversion, status quo bias, and other factors. Further, we are able to identify and validate variation across demographic subgroups. In the supporting information (p. 14), we provide further discussion of the lessons we draw for future experimental design, including discussion of trade-offs associated with providing more complicated sets of changes to respondents, varying presentation of tax costs by respondent income level, and specification of the change proposal distribution.

Applications

Our approach translates readily to applications other than the British national budget, with appropriate domestication of the categories. In the United States, a

---

[10]Our preregistration documents (Barnes, Blumenau, and Lauderdale 2019) can be found at http://egap.org/registration/5550.

FIGURE 6 Levels of Support for Validation Study Proposals
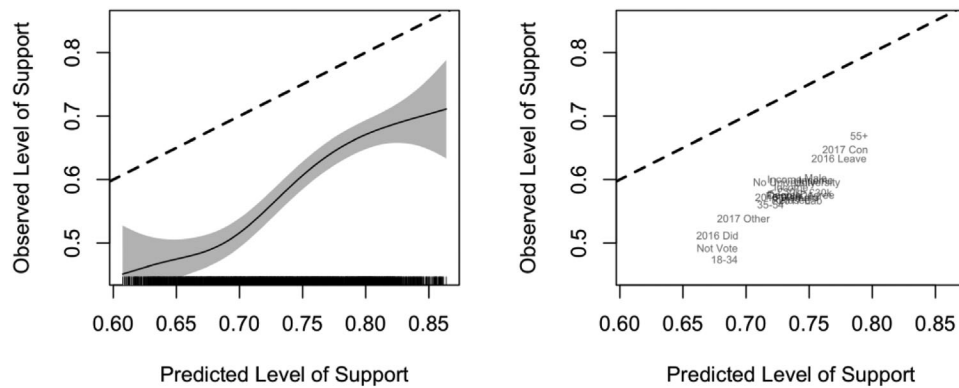
Note: The left panel displays observed levels of support for proposals in the validation study as a function of predicted level of support given the set of explanatory variables in the multivariate model. Perfect agreement between the predictions and the validation would correspond to the dashed line; the black line is a spline fit to the validation data. The right panel shows the observed level of support as a function of predicted level of support for different groups.

parallel implementation might use categories from the president's published budget, as in Bonica (2015). There is no general right answer regarding which categories to use, but the choice ought to be as accessible as possible to a typical citizen. First, the measurement strategy outlined here could be usefully incorporated into existing large-scale survey projects in order to measure these preferences in a quantitatively meaningful way over time and across space. Building this strategy into national election studies would provide a consistent way to track relative spending priorities, and the relative priority of tax reductions, within countries over time. One natural cross-national application would be to use the budget categories of the European Union, and to use our design to measure average preferences across these areas in different EU member states. Our approach can also provide a way to gauge preferences over spending trade-offs *within* policy areas. For example, within the area of education, voters may prefer allocating more resources to early-years intervention at the cost of lower K–12 spending.

Second, researchers can use this strategy to measure spending preference outcomes in bespoke experiments, including—and perhaps especially—in the context of studies focused on one area of social policy spending. One problem that such studies face when asking about budgetary preferences is that questions about spending in that area typically follow a series of other questions about that spending area, creating potential for both framing and demand effects. Our approach has similar advantages to the use of conjoint experiments as a way to minimize social desirability bias.

Third, we have not exhausted the kinds of analysis that can be done with the data we have collected already. For example, because our estimation strategy enables estimation of a linear model for preferences in terms of covariates, the derived preferences could be used to generate descriptions of how spending preferences are geographically distributed via poststratification. This would generate geographic estimates of public budget preferences, which might be useful for assessing campaigns, intraparty variation in budget preferences of representatives, and other research questions.

Fourth, estimating the public preference distribution enables a range of welfare and social choice analyses. The first kind of welfare analysis that we might do would assess the aggregate social welfare deriving from different budgets. A second kind of welfare analysis, facilitated by our covariate model, might assess which population groups are more or less satisfied with the status quo or particular proposals. In addition to welfare questions, we can also ask social choice questions about the estimated preference distribution. A first type of social choice analysis would be to assess whether particular spending profiles would defeat the status quo spending profile. Such analyses can be extended to characterize relevant regions of the winset: all proposals that would defeat the status quo by majority rule. A second type of social choice analysis involves identifying extreme budgets subject to popularity constraints, such as the biggest overall tax/spending cut or increase that defeats the status quo. A third type of social choice analysis involves identifying popularity-maximizing budgets subject to total spending constraints. In the supporting information

(p. 15), we discuss each of these welfare and social choice analyses in more detail and provide examples of two of them.

Conclusions

We have demonstrated a new strategy to measure spending preferences multidimensionally and quantitatively. We elicit quantitative preferences on narrow spending trade-offs from a sample of respondents who are representative of the broader public on measurable characteristics and then translate those into estimates of a distribution of public spending preferences. Doing this involves assuming coherent preferences, but we have argued this is a sensible approach because it enables us to find a distribution of coherent budget preferences that is consistent with the choices that citizens make when faced with accessible trade-offs between spending categories and taxation. Whether or not individual members of the public in fact hold or can articulate consistent budgetary preferences, it is worth knowing which feasible budgets are likely to best match citizens' assessments of the relevant tax and spending trade-offs.

Our approach to the study of spending preferences serves as an example of three broader methodological principles that are widely applicable. First, and most narrowly, our approach highlights the utility of using continuous treatments within multivariate choice experiments, of which conjoint experiments are one prominent example. Since the central tool for analyzing the data generated in conjoint experiments is regression analysis, and regression analysis is suited for the analyses of continuous variables, the focus on qualitative attribute levels in conjoint approaches often represents a missed opportunity to generate more precisely quantifiable effects. Second, structural models are a useful tool for estimating aggregate quantities of interest in contexts where we have insufficient data to generate complete descriptions of the distributions at the individual level and where regression-based approaches are either inapplicable or generate relatively uninteresting causal estimands. Third, this article emphasizes the importance of validation. In any approach based on a parametric model, we should be concerned about the degree to which the assumptions made about functional forms drive the results. The validation exercise not only provides greater confidence in the results obtained, but also highlights specific areas for further development.

Substantively, we find that current levels of government spending in the United Kingndom are lower than the average preferences of the UK public, and that this is true across many different spending categories. Crucially, as the setup of our experiment forces respondents to take government budget constraints seriously, the public expresses these preferences for increased spending even when taking into account increases in taxation required to fund them. We showed that a public budget including an implied tax hike of approximately 7% received majority support among those expressing an opinion. We are also able to identify demographic variation in preferences for more spending that operates in different directions in different spending areas. Our finding that it is older voters, not younger voters, who want to substantially increase UK government spending is clear in the results of both the original experiment and the validation study and is at odds with the conventional wisdom regarding age and political preferences in the United Kingdom. The fact that this age pattern is present both unconditionally and also conditional on gender, education, income, EU referendum vote, and general election vote suggests this is likely to be a real and important feature of contemporary British politics that merits further investigation.

References

Arceneaux, Kevin. 2002. "Direct Democracy and the Link between Public Opinion and State Abortion Policy." *State Politics & Policy Quarterly* 2(4): 372–87.

Ballard-Rosa, Cameron, Lucy Martin, and Kenneth Scheve. 2017. "The Structure of American Income Tax Policy Preferences." *Journal of Politics* 79(1): 1–16. https://doi.org/10.1086/687324.

Barnes, Lucy, Jack Blumenau, and Benjamin E. Lauderdale. 2019. "Measuring UK Attitudes towards Public Spending: Validation of Predictions from a Multivariate Tax Summary Experiment." EGAP ID: 20190301AA. http://egap.org/registration/5550.

Bartels, Larry M. 2005. "Homer Gets a Tax Cut: Inequality and Public Policy in the American Mind." *Perspectives on Politics* 3(1): 15–31.

Bechtel, Michael M., Federica Genovese, and Kenneth F. Scheve. 2017. "Interests, Norms and Support for the Provision of Global Public Goods: The Case of Climate Cooperation." *British Journal of Political Science* 49(4): 1333–55.

Berinsky, Adam J. 2008. "Survey Non-Response." In *The SAGE Handbook of Public Opinion Research* ed. Wolfgang Donsbach and Michael W. Traugott. London: Sage, 309–21.

Bonica, Adam. 2015. "Measuring Public Spending Preferences Using an Interactive Budgeting Questionnaire." *Research & Politics* 2(2). https://doi.org/10.1177/2053168015586471.

Branham, J. Alexander, and Stephen A. Jessee. 2017. "Modeling Spending Preferences & Public Policy." *Electoral Studies* 49: 155–72.

Bremer, Björn, and Reto Bürgisser. 2019. "Attitudes towards Fiscal Consolidation: Evidence from Survey Experiments in Four European Countries." http://www.bjoern-bremer.com/projects/trade-offs.

Broockman, David E. 2016. "Approaches to Studying Policy Representation." *Legislative Studies Quarterly* 41(1): 181–215.

Busemeyer, Marius R., Julian L. Garritzmann, Erik Neimanns, and Roula Nezi. 2018. "Investing in Education in Europe: Evidence from a New Survey of Public Opinion." *Journal of European Social Policy* 28(1): 34–54.

Carpenter, Bob, Andrew Gelman, Matt Hoffman, Daniel Lee, Ben Goodrich, Michael Betancourt, Michael A. Brubaker, Jiqiang Guo, Peter Li, and Allen Riddell. 2016. "Stan: A Probabilistic Programming Language." *Journal of Statistical Software* 76(1): 1–32.

Caughey, Devin, Tom O'Grady, and Christopher Warshaw. 2019. "Policy Ideology in European Mass Publics, 1981–2016." *American Political Science Review* 113(3): 674–93.

Cavaillé, Charlotte, and Kris-Stella Trump. 2015. "The Two Facets of Social Policy Preferences." *Journal of Politics* 77(1): 146–60.

Citrin, Jack. 1979. "Do People Want Something for Nothing? Public Opinion on Taxes and Government Spending." *National Tax Journal* 32(2): 113–29.

Clinton, Joshua, Simon Jackman, and Douglas Rivers. 2004. "The Statistical Analysis of Roll Call Data." *American Political Science Review* 98(2): 355–70.

D'Attoma, John, Kim-Lee Tuxhorn, and Sven Steinmo. 2018. "Measuring Budgetary Preferences Using Interactive Budget Simulations: A Holistic Approach." https://www.researchgate.net/publication/324163964_A_Holistic_Approach_to_Studying_Budget_Preferences_Using_Interactive_Budget_Tools_for_Social_Science_Research

Downs, Anthony. 1960. "Why the Government Budget Is Too Small in a Democracy." *World Politics* 12(4): 541–63.

Edlund, Jonas, and Ingemar Johansson Sevä. 2013. "Exploring the 'Something for Nothing' Syndrome: Confused Citizens or Free Riders? Evidence from Sweden." *Scandinavian Political Studies* 36(4): 293–319.

Gilens, Martin. 2012. *Affluence and Influence: Economic Inequality and Political Power in America*. Princeton, NJ: Princeton University Press.

Grasso, Maria Teresa, Stephen Farrall, Emily Gray, Colin Hay, and Will Jennings. 2019. "Thatcher's Children, Blair's Babies, Political Socialization and Trickle-Down Value Change: An Age, Period and Cohort Analysis." *British Journal of Political Science* 49(1): 17–36.

Hainmueller, Jens, and Daniel J. Hopkins. 2015. "The Hidden American Immigration Consensus: A Conjoint Analysis of Attitudes toward Immigrants." *American Journal of Political Science* 59(3): 529–48.

Hainmueller, Jens, Daniel J. Hopkins, and Teppei Yamamoto. 2014. "Causal Inference in Conjoint Analysis: Understanding Multi-Dimensional Choices via Stated Preferences Experiments." *Political Analysis* 22(1): 1–30.

Hansen, John Mark. 1998. "Individuals, Institutions, and Public Preferences over Public Finance." *American Political Science Review* 92(3): 513–31.

Häusermann, Silja. 2010. *The Politics of Welfare State Reform in Continental Europe: Modernization in Hard Times*. Cambridge: Cambridge University Press.

Hübscher, Evelyne, Thomas Sattler, and Zbigniew Truchlewski. 2020. "Voter Attitudes towards Fiscal Trade-Offs: Evidence from Three European Countries." http://www.thomassattler.org/mediapool/128/1289817/data/HST2020.pdf.

Hübscher Evelyne, Thomas Sattler, and Markus Wagner. 2020. "Voter Responses to Fiscal Austerity." *British Journal of Political Science*. https://doi.org/10.1017/S0007123420000320.

Kastellec, Jonathan P., Jeffrey R. Lax, and Justin H. Phillips. 2010. "Public Opinion and Senate Confirmation of Supreme Court Nominees." *Journal of Politics* 72(3): 767–84.

Kelly, Nathan J., and Peter K. Enns. 2010. "Inequality and the Dynamics of Public Opinion: The Self-Reinforcing Link between Economic Inequality and Mass Preferences." *American Journal of Political Science* 54(4): 855–70.

Lauderdale, Benjamin E., Chris Hanretty, and Nick Vivyan. 2018. "Decomposing Public Opinion Variation into Ideology, Idiosyncrasy, and Instability." *Journal of Politics* 80(2): 707–12.

Lax, Jeffrey R., and Justin H. Phillips. 2009a. "Gay Rights in the States: Public Opinion and Policy Responsiveness." *American Political Science Review* 103(3): 367–86.

Lax, Jeffrey R., and Justin H. Phillips. 2009b. "How Should We Estimate Public Opinion in the States?" *American Journal of Political Science* 53(1): 107–21.

McCall, Leslie, and Lane Kenworthy. 2009. "Americans' Social Policy Preferences in the Era of Rising Inequality." *Perspectives on Politics* 7(3): 459–84.

Meltzer, Allan H., and Scott F. Richard. 1981. "A Rational Theory of the Size of Government." *Journal of Political Economy* 89(5): 914–27.

Mettler, Suzanne. 2011. *The Submerged State: How Invisible Government Policies Undermine American Democracy*. Chicago: University of Chicago Press.

O'Grady, Tom. 2020. "Is Europe Becoming a Gerontocracy? New Evidence on Age Cleavages in Europe since the 1980s." Working Paper, University College London. https://osf.io/preprints/socarxiv/r8xbg/.

Page, Benjamin I., and Lawrence R. Jacobs. 2009. Class War? What Americans Really *Think about Economic Inequality*. Chicago: University of Chicago Press.

Rehm, Philipp. 2011. "Social Policy by Popular Demand." *World Politics* 63(2): 271–99.

Rueda, David, and Daniel Stegmueller. 2019. *Who Wants What? Redistribution Preferences in Comparative Perspective*. Cambridge Studies in Comparative Politics. Cambridge: Cambridge University Press.

Soroka, Stuart N., and Christopher Wlezien. 2005. "Opinion–Policy Dynamics: Public Preferences and Public

Expenditure in the United Kingdom." *British Journal of Political Science* 35(4): 665–89.

Soroka, Stuart N., and Christopher Wlezien. 2009. *Degrees of Democracy: Politics, Public Opinion, and Policy.* Cambridge: Cambridge University Press.

Svallfors, Stefan. 2012. "Welfare States and Welfare Attitudes." In *Contested Welfare States: Welfare Attitudes in Europe and Beyond*, ed. Stefan Svallfors. Stanford, CA: Stanford University Press: 1–24.

Williamson, Vanessa S. 2017. *Read My Lips: Why Americans Are Proud to Pay Taxes.* Princeton, NJ: Princeton University Press.

Winter, Soren, and Poul Erik Mouritzen. 2001. "Why People Want Something for Nothing: The Role of Asymmetrical Illusions." *European Journal of Political Research* 39(1): 109–43.

Wlezien, Christopher 1995. "The Public as Thermostat: Dynamics of Preferences for Spending." *American Journal of Political Science* 39(4): 981–1000.

Supporting Information

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

Appendix A: Tax Summary and Experiment Screenshots
Appendix B: Components of Tax Summary Spending Areas
Appendix C: Proposal Distribution
Appendix D: Further Discussion of Choice Model Assumptions
Appendix E: Incidental Model Parameters
Appendix F: Estimates Under Alternative Model Specifications and Data Subsets
Appendix G: Principal Components Analysis on Estimated Preference Distributions
Appendix H: Lessons for Experimental Design
Appendix I: Welfare and Social Choice Analysis