# How Many Americans Hold Policy Preferences that are Related Across Issues?

Matthew Tyler[†] William Marble[‡]

Draft Version: December 6, 2018

**Abstract**

To what extent do Americans hold policy opinions that "go together" ideologically? On one account, most Americans have seemingly independent views across issues. On another account, most Americans hold opinions that have a common structure — a structure that ideal point models can reveal. We use cross-validation and a new ideal point estimator to study how predictable preferences are at an individual level. Applying these methods to survey data of the public and politicians, we show that the vast majority of Americans provide survey responses that are more predictable with ideal point models (unidimensional and multidimensional) than without. This increased predictability ranges from a high average increase among politicians to only a slight average increase among Americans who cannot place the parties on the correct side of the conventional left-right spectrum. Taken together, our results suggest that most Americans do hold opinions that follow a common structure, but how much this structure explains opinions varies significantly across individuals.

**Word Count: 777**

[†]PhD Candidate, Department of Political Science, Stanford University. mdtyler@stanford.edu
[‡]PhD Candidate, Department of Political Science, Stanford University. wpmarble@stanford.edu

# 1 Introduction

A classic literature in political science argues that, because correlations between policy preferences across issues are low, American policy preferences are mostly independent across issues (Converse, 1964). Under this view, only a small minority of the most politically-sophisticated Americans have policy preferences that are constrained across issues. More recent evidence has confirmed that across-issue correlations remain low, even in era of polarization (Sullivan, Piereson and Marcus, 1978; Delli Carpini and Keeter, 1997; Kinder, 2003; Broockman, 2016).

In contrast, another segment of the literature starts from the assumption that most Americans have an ideological structure for their policy preferences, but surveyed policy preferences measure this ideological structure with a great amount of error (Achen, 1975; Ansolabehere, Rodden and Snyder, 2008). The most recent iterations of this view take the form of ideal point estimators (e.g., Treier and Hillygus, 2009; Jessee, 2009; Bafumi and Herron, 2010; Tausanovitch and Warshaw, 2013; Hill and Tausanovitch, 2015). Defenders of this approach argue that there is a strong correlation between ideal point estimates and voting behavior (Ansolabehere, Rodden and Snyder, 2008; Jessee, 2009).

Clearly these two stories can not apply equally well to most Americans. Furthermore, correlations between policy questions or between ideal points and vote choice do not tell us how many or which Americans these models fit better. The objective of this paper is to (1) quantify how many Americans ideal point models fit better than a null model of independent issues and (2) predict *which* Americans have their policy preferences explained better by ideal point models.

First, we clarify why the existing work has failed to answer these questions: there is as yet no agreed-upon measure of whether an individual's preferences are related across issues. Correlation-based measures are necessarily aggregate measures that hide individual heterogeneity. Moreover, existing measures make it difficult to quantify *how many* Americans hold attitudes consistent with an ideal point model.

Next, to resolve this empirical deadlock, we develop an individual-level measure of how well ideal point models fit compared to the independent preference model. Our measure is the log likelihood ratio of the two competing models, which is highly sensitive to the relative predictive power of each model. We show how to use this measure to infer how many and which Americans are better modeled by ideal points than independent preferences.

We find that ideal point models explain survey responses better than an indepednent preference model for the vast majority (> 90%) of Americans. This majority further increases when we use multidimensional rather than unidimensional ideal point models, implying that a nontrivial amount of American policy preferences are structured in a multidimensional way. These results suggest that claims that most Americans do not have "meaningful" attitudes are unfounded.

However, we also find that there is significant heterogeneity in how much better ideal point models describe attitudes than an independent-preference model. We use our individual-level measure to explore which covariates predict an individual having survey responses more consistent with an ideal point model. Not surprisingly, we find that politicians' attitudes are extremely well-explained by ideal point models. In the mass public, survey respondents who can correctly place candidates on the correct side of the left-right spectrum, sorted partisans, political donors, and Republicans receive the greatest predictability boost from ideal point models. These correlations hold in both bivariate and multivariate analyses.

Finally, in addition to introducing the log-likelihood ratio as a measure of how well a survey respondents' attitudes "go together" and using this measure to estimate how many and which Americans have related attituedes, we also make two additional methodological contributions to the literature on ideal points.

First, because we are interested in computing model fit statistics, we must deal with the possibility of overfitting. In order to generate unbiased estimates of model fit, we cannot use the same data that was used to estimate the model. To overcome this problem, we introduce a cross-validation technique that allows us to characterize the likelihood of a vector of individual-level survey responses, given a model estimate.

Second, our individual-level measure of ideal point fit requires us to know the likelihoods of different policy preferences under unidimensional and multidimensional ideal point models, and the cross-validation procedure requires re-estimating these models many times. We apply a recently developed technique in machine learning for fitting continuous latent variable models to ideal point models, allowing us to quickly estimate multidimensional ideal point models with the categorical responses typically found in surveys. For our purposes, this method appears to outperform the state-of-the-art techniques used in political science and psychometrics for fitting latent variable models, and contributes to a growing literature on fast estimation of complex ideal point models (e.g., Imai, Lo and Olmsted, 2016; Goplerud, 2018).

# 2 Theoretical Framework: Independent Preferences versus Ideal Points

In this section we set the stage for determining how many and which Americans hold policy preferences that are related across issues in a common structure. After first reviewing the extensive literature on this topic, we identify why the existing literature has so far been unable to provide an answer.

## 2.1 Literature

Since the advent of modern public opinion research, scholars have debated whether Americans hold political opinions that "go together" ideologically. Seminal research by Converse (1964) argues that the average citizen's conception of politics bore little relation to the ideological structure that characterized politicians' and political elites' preferences. Converse claimed that "large portions of an electorate ... simply do not have meaningful beliefs, even on issues that have formed the basis for intense political controversy among elites for substantial periods of time" (245). On this account, Americans political opinions tend to be unrelated to each other, devoid of a unifying framework.

The primary mode of testing this argument is measuring the correlations between survey respondents' opinions across a range of policy issues. A high degree of correlation indicates that respondents' beliefs are "constrained" — taken to be evidence that these beliefs are governed by an underlying disposition toward politics. In the case of low correlations, beliefs are unconstrained — that is, they are not governed by an underlying philosophy. Empirically, interitem correlations across policy questions tend to be quite low, generally lower than 0.25 (e.g., Converse 1964, 228; Broockman 2016, 193). The conclusion — reinforced by a large subsequent literature (e.g., Delli Carpini and Keeter, 1997; Sullivan, Piereson and Marcus, 1978; Kinder, 2003) — is that citizens' political views across issues tend to be unrelated. One influential explanation, offered by Zaller and Feldman (1992), is that citizens hold multiple, conflicting principles about politics, and when answering survey questions, they sample only a few to inform their response. If their sampling process leads them to consider different principles when answering different questions, their opinions will seem inconsistent, yielding low interitem correlations.

On the other hand, other researchers suggest that the low interitem correlations are not due to the underlying preferences being unrelated, but rather due to measurement error in surveys (Achen, 1975; Ansolabehere, Rodden and Snyder, 2008). On this account, surveys are an imperfect tools that introduce noise into measure of political attitudes, thereby depressing correlations. In this case, it is possible to uncover preference structures from survey data using appropriate methods — typically ideal point models — to account for measurement error. A large literature has emerged that applies ideal point models to survey data (e.g., Hill and Tausanovitch, 2015; Tausanovitch and Warshaw, 2013; Treier and Hillygus, 2009; Jessee, 2009; Bafumi and Herron, 2010; Pan and Xu, 2017). The core assumption of this work is that we can extract a politically meaningful measure of ideology by examining the joint distribution of survey responses.

However, these two views — the Conversian argument that opinions across issues are unrelated and the ideal point view that survey responses can reveal a low-dimensional structure — are mutually incompatible at the individual level. However, there is no reason that either theory must hold for every member of the public. The core question, as we see it, is how many people have independent-preferences and how many people have related preferences? A different understanding of public opinion would emerge if, for example, each theory successfully described half of the population. A secondary yet substantively important question is, for which members of the public does an ideal point model outperform an independent-response model? The literature has not adequately answered these questions for three reasons.[1]

First, there are divergent standards of evidence amongst the two sets of researchers. As Freeder, Lenz and Turney (2018) note, the literature inspired by Converse focuses on interitem correlations due to "tradition." The other group of researchers, on the other hand, typically point to the correlation between ideal points and voting behavior or increases in in-sample prediction of survey responses as a justification for using ideal point models.

Second, the debate has been typically framed in black-and-white terms over whether citizens writ large have "meaningful" preferences. The qualitative nature of the debate has obscured nuance. No model will be able to perfectly predict survey responses, so it is more useful to quantify the extent to which one model fits better than another rather than declaring that citizens do or do not have meaningful attitudes. Additionally, any given model will apply better to some members of the public than others. To our knowledge, there have been no attempts to estimate, at an individual level, the proportion of the population that holds opinions better explained by an ideal point model than a model in which opinions are unrelated. Further, the measurement error hypothesis typically treats the entire population as suffering from the same level of measurement error in surveys, and therefore cannot investigate variance in how well measurement-error corrections perform for different subsets of the population.[2]

[1] For clarity, we note that we pose these questions purely in terms of cross-sectional associations between policy items. Much research has focused on whether opinions *within* a given issue are stable over time; this question is related to but distinct from the question of whether attitudes are related *across* issues.

[2] Some scholars allow the measurement error of survey responses to vary at the individual level (e.g., **?**), but this is still insufficient for identifying how many individuals have related preferences.

Third, and relatedly, different segments of the population may have structured preferences on different sets of issues. For example, one could imagine that hard-line moral conservatives would be consistent on one set of issues, and libertarian conservatives would be consistent on another, non-overlapping, set of issues. The existence of such distinct issue publics would serve to depress interitem correlations in the population, yet more flexible models would be able to detect such complex dependency structures.

Below, we show how to overcome these obstacles. We estimate both independent preference models and ideal point models, then evaluate them using out-of-sample fit statistics. We can then estimate the proportion of the population that is better characterized by each model. This approach also allows us to characterize, at an individual level, which model better explains survey responses.

Our approach is most closely related to four papers. Lauderdale, Hanretty and Vivyan (2017) use an ideal point model to decompose the variation in survey responses into three components: ideology, by which they mean a preference structure common to all respondents; idiosyncrasy, by which they mean persistent individual-level deviations from the common structure; and instability, by which they mean response instability that is induced by the measurement technique. They find that roughly 15% of opinion variation can be explained by ideology. We similarly seek to characterize the extent to which ideal point models can explain survey responses, but our goal of estimating the proportion of Americans ideal point models apply to is fundamentally different, we employ an out-of-sample validation technique, and examine a larger array of data sources.

Freeder, Lenz and Turney (2018) examine the Converse critique from a different angle. Instead of measuring interitem correlations, they examine test-retest reliability as a measure of "meaningful" preferences. They show that people who can correctly place candidates on a left-right spectrum across a range of issues exhibit higher stability in survey responses. They conclude that this group of people, which constitutes about 25-50% of the public, is likely to have meaningful preferences.[3] Similarly, Barber and Pope (2017) argue that issue consistency — that is, the correlation of opinions across issues — is a function of political knowledge, with high-knowledge respondents exhibiting higher issue consistency. Following this work, below we examine the extent to which knowledge of "what goes with what" is correlated with increased performance of ideal point models relative to a null model.

Finally, Hill and Kriesi (2001) also attack the problem of over-time stability in preferences. They estimate a finite mixture model in which there are three groups — people with stable opinions, people whose opinions change over time, and people who do not have genuine opinions. Using Swiss panel data, they find that over half of the public have stable opinions, and under 10% have no genuine opinions. We also seek to characterize the proportion of the population that falls into certain groups using a mixture of "null" and "alternative" types in the population. However, we focus on the structure of opinions across issues, rather than across time, and introduce new methods to estimate the proportions.

## 3 A New Measure of Related Preferences

In this section we try to resolve the empirical deadlock suggested in section 2 by developing a measure of how likely an individual's preferences are to have emerged from an ideal point model with common structure versus a model that assumes issue preferences are independent. We then describe a method for using this measure to infer how many and which Americans ideal point models accurately describe.

[3] See also Sniderman and Stiglitz (2012), who argue that knowing how candidates and parties line up across issues is an important predictor of attitude structures.

| | | Healthcare | |
|---|---|---|---|
| | | C | D |
| **Taxes** | A | 0.25 | 0.25 |
| | B | 0.25 | 0.25 |

(a) Independent Preferences

| | | Healthcare | |
|---|---|---|---|
| | | C | D |
| **Taxes** | A | 0.35 | 0.15 |
| | B | 0.15 | 0.35 |

(b) Preferences Implied by an Ideal Point Model

Table 1: Example preference distributions where cell entries represent probability of an individual picking that policy pair (AC, AD, BC, or BD). Certain response patterns are much more likely to occur under one of the two distributions.

## 3.1 Using Log-Likelihood Ratios to Compare Independent and Ideal Point Models

Our measure of ideal point fit is based on calculating the likelihood of an individual's preferences under two different probability distributions: one distribution that assumes independent preferences and another that assumes preferences are related according to an ideal point model. The proposed measure attains a higher value when an individual provides policy preferences that are more consistent with an ideal point model than independent issues.

To see how we calculate the measure in a small-scale example, and how we can use these likelihoods to infer the appropriateness of ideal point models, consider the crosstabs given in Tables 1a-1b. Each Table is a distribution of stated policy preferences for imaginary tax policies A, B and imaginary healthcare policies C, D. Every individual states pairs of preferences: AC, AD, BC, or BD. Table 1a represents a distribution of independent policy preferences. We know the preferences are independent because the joint probabilities factor; e.g., $P(AC) = P(A)P(C)$. Table 1b is a distribution where the probabilities of preference pairs is implied by a particular ideal point model.[4] Note that the preferences are dependent; e.g., $P(AC) \neq P(A)P(C)$. The ideal point model induces a correlation between certain types of tax and healthcare policies

[4]Table 1b was generated by assuming that, given an individual's $z \sim \mathcal{N}(0,1)$, policy choices are conditionally independent with the probability of Taxes = A being $\Lambda(\beta_T z)$ and the probability of Healthcare = C being $\Lambda(\beta_H z)$, where $\beta_T = \beta_H = 2$ and $\Lambda(u) = 1/(1 + e^{-u})$ is the standard logistic function. We say this distribution has a common structure because every individual's latent variable $z$ is linked to policy preferences through the common parameters $\beta_T, \beta_H$.

because it makes certain choice pairs like AC and BD much more likely to be observed relative to choices like AD and BC. This is often because AC and BD are somehow logically connected policies, but from a modeling perspective this is not strictly necessary. We will denote the probability distribution given in Table 1a with $P_a$ and the probability distribution given in Table 1b with $P_b$.

The core idea behind our measure of ideal point model suitability is that when an individual's policy preference $x$ is drawn from an ideal point model like that implied by Table 1b (i.e., $x \sim P_b$) we should expect $P_b(x) > P_a(x)$, and when $x$ is drawn according to a independent distribution like that given in Table 1a (i.e., $x \sim P_a$) we should expect $P_a(x) > P_b(x)$. Formally, define the LLR statistic as the log of the likelihood ratio:

$$\Delta(x) = \log \frac{P_b(x)}{P_a(x)} = \log P_b(x) - \log P_a(x). \tag{1}$$

Rather than the log of the likelihood ratio, we can also think of it as the change in log likelihood going from the independent model to the ideal point model. It is immediate that $\Delta(x)$ is positive when $x$ is more likely under the ideal point model than under the independent model, and negative if the opposite is true. For example, AC has a higher probability under the ideal point model and $\Delta(\text{AC}) \approx 0.34$. Similarly, AD has a higher probability under the independent model and $\Delta(\text{AD}) \approx -0.51$. Without much trouble, one can show that for data drawn from the ideal point model we have $\Delta(x) > 0$ on average: $E_{P_b}[\Delta(x)] \geq 0$, and similarly if our data is drawn from an independent model then we have $\Delta(x) < 0$ on average: $E_{P_a}[\Delta(x)] \leq 0$, where $E_P$ is the expectation taken with respect to $P$.[5]

We can generalize the proposed measure to a more general context with multiple individuals, issues, different types of response options, and multiple ideal models. Suppose for each individual $i = 1, \dots, N$ we observe a set of $J$ policy preferences $x_i = (x_{i1}, x_{i2}, \dots, x_{iJ})$. Suppose there are two different types of models under consideration: an independent preference

[5] Some readers will notice that $E_{P_b}[\Delta(x)]$ is the Kullback-Leibler (KL) divergence between $P_b$ and $P_a$ while $E_{P_a}[\Delta(x)]$ is the negative KL divergence between $P_a$ and $P_b$.

model $P_0$ and $D$ ideal point models models $P_1, P_2, \ldots, P_D$. In this context, $D$ different models correspond to ideal point models with $d = 1, \ldots, D$ dimensions (e.g., for $d = 2$ we might expect an economic and moral dimension, Ansolabehere, Rodden and Snyder, 2006). Define the log-likelihood ratio (LLR) of model $d$ for individual $i$ as

$$\Delta_d(x_i) = \log \frac{P_d(x_i)}{P_0(x_i)} = \log P_d(x_i) - \log P_0(x_i) \quad \text{for each } d = 1, \ldots, D. \tag{2}$$

This measure has similar properties as when there were only two distributions under consideration; namely, that $E_{P_d}[\Delta_d(x_i)] > 0$ and $E_{P_0}[\Delta_d(x_i)] < 0$ for each $d = 1, \ldots, D$.

As we argue below, the LLR improves on earlier measures because it can answer questions like "how many" and "which" individuals have preferences that ideal points explain better, but that is not all: the LLR has a useful statistical motivation as well. The LLR has the desirable property that for testing the null hypothesis that $x_i \sim P_0$ versus the alternative hypothesis $x_i \sim P_d$, tests of the form "reject if $\Delta_d(x_i) > c$" for some threshold $c$ are the most powerful tests at their significance level.[6] For our purposes (we will not be doing hypothesis tests), this result is suggestive that the LLR statistic is highly likely to notice when the ideal point model $P_d$ is better at explaining the data point $x_i$ than the independent model $P_0$.

## 3.2 Using LLRs to Estimate How Many and Which Americans are Described Better by Ideal Point Models than Independent Preferences

Having proposed the LLR measure, we now show how to use it to estimate how many Americans have preferences are better described by ideal point models than independent preferences. Taking some cues from the large-scale hypothesis testing literature in statistics (such as Efron, 2004), we show that this problem can be reduced to estimating the weight in a two-component mixture model.

To estimate the proportion of ideal point types, we can imagine that there are two types of

[6]This is known as the Neyman-Pearson Lemma (see section 3.2 of Lehmann and Romano, 2006). Determining the right threshold $c$ would depend on $P_0$ and our desired significance level.

individuals. The independent types are individuals whose preferences are justifiably modeled as independent; that is, $x_i \sim P_0$ where $P_0$ is an independent preference model. For ideal point types, we assume their policy preferences are drawn according to an ideal point model, so that $x_i \sim P_d$ for some dimensionality $d \in \{1,\ldots,D\}$. Denote the proportion of the second type of individual as $\pi_d \in [0,1]$. For now, we assume both of these distributions are known and postpone explaining how we estimate them until section 4.

With this setup, we have reduced the the task of estimating how many Americans are better explained by ideals point to the problem of estimating the parameter $\pi_d$. Introducing some more notation, let $F_0$ be the distribution of $\Delta_d(x_i)$ when $x_i \sim P_0$, and similarly define $F_d$ as the distribution of $\Delta_d(x_i)$ when $x_i \sim P_d$. These distributions capture the fact that the LLR statistic will have a different distribution under the “null” hypothesis that preferences are independent than under the “alternative” hypothesis that preferences are drawn according to an ideal point model. Figure 1 illustrates how different these distributions can be on data from a survey of sitting state legislators (Broockman, 2016). Since we do not know which individuals belong to which group, then it must be the case that the marginal distribution of the LLR is a $\pi_d$-weighted average of $F_d$ and $F_0$. Formally,

$$\Delta_d(x_i) \sim \pi_d F_d + (1-\pi_d)F_0. \tag{3}$$

This is useful because it means we can write the average log likelihood function of $\pi_d$ as

$$\frac{1}{N}\sum_{i=1}^{N}\log\Big[\pi_d F_d(\Delta_d(x_i)) + (1-\pi_d)F_0(\Delta_d(x_i))\Big]. \tag{4}$$

When we maximize this, we find the value $\hat{\pi}_d$ that makes the empirical distribution of our LLR statistic as close as possible to the $\hat{\pi}_d$ weighted average of the ideal point distribution $F_d$ and the independent distribution $F_0$. We use $\hat{\pi}_d$ to estimate $\pi_d$.

This approach offers us an answer to our main question of *how many* Americans have related preferences (based on an ideal point model) without having to perform hypothesis tests

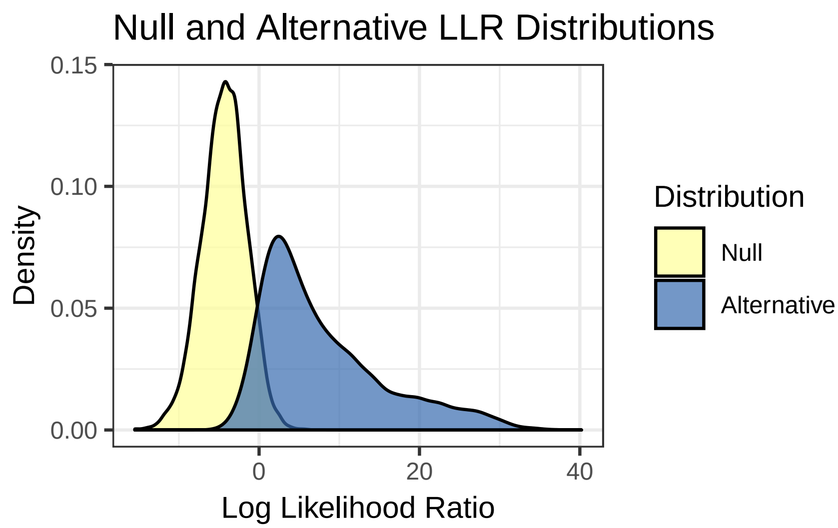



Figure 1: Holdout log likelihood ratio (LLR) associated with a 1-dimensional ideal point model relative to an unrelated-preference model . The input data are 31 survey questions on answered by sitting state legislators, collected by Broockman (2016) The yellow density estimate shows the LLR distribution when we assume survey responses are drawn from a null distribution of unrelated preferences, while the blue density estimate shows the LLR distribution when we assume survey responses are drawn from a unidimensional ideal point model. LLRs generated from the null distribution are usually negative, and LLRs generated from the alternative distribution are usually positive, but there is non-negligible overlap even for these politicians.

on each individual. Hypothesis tests would be far too conservative in this context since each individual is only observed in our data once. Even if we applied more liberal testing procedures such as controlling for false-discover rates (FDR, see Benjamini and Hochberg, 1995), almost all individuals would be rejected.[7] However, since our primary goal is not to rule whether any particular individual $i$ has related preferences, but rather to estimate the frequency of individuals with related preferences, it is appropriate to focus on just estimating the mixture weight $\pi_d$.

One could argue that instead of using the marginal distribution of $\Delta_d(x_i)$ to identify $\pi_d$, we could just use the marginal distribution $x_i$ and estimate $\pi_d$ from a mixture model (i.e., model $x_i \sim \pi_d P_d + (1-\pi_d)P_0$). However, we do not like this approach as much. Our primary objection

[7] For instance, when we apply the FDR control method to a sample of state legislators and members of Congress from the early 2010s, we rejected the null hypothesis of independent preferences for less than fifty percent of politicians. This seems far too conservative, since it is commonly taken for granted that almost all politicians subscribe to a common elite ideological structure and can be modeled as having ideal points. In contrast, our estimator $\hat{\pi}_d$ described above finds that every politician in our sample is better modeled by using ideal points than independent preferences.

is that, unlike $x_i$, the LLR is a univariate statistic whose distribution can be visually inspected (as seen in Figure 1). By looking at the distributions $F_0, F_d$ and the marginal distribution of $\Delta_d(x_i)$ in our data, we can (1) confirm which of $F_0, F_d$ a priori looks like a better fit to the marginal distribution of $\Delta_d(x_i)$ and (2) confirm whether the distribution implied by the estimate $\hat{\pi}_d$ is a resonable approximation of the marginal distribution of $\Delta_d(x_i)$. Second, the approach we outlined above is more consistent with the statistic mixture-model approach used in large-scale testing where the results of thousands of multivariate hypothesis are often reduced to one statistic, such as a z-value, corresponding to our $\Delta_d(x_i)$ (e.g., Efron, 2004). Finally, as described in section 4.3, we use cross-validation when evaluating the model likelihoods to avoid overfitting concerns, which precludes us from simultaneously estimating $\pi_d, P_0, P_d$ in a traditional mixture model estimator. With all that said, appendix D shows the two mixture-model approaches of finding $\pi_d$ (using cross-validation) offer extremely similar results and our main conclusions are robust to how we construct the mixture model.

In addition to estimating how many Americans are better explained by ideal point models, we can also use the individual LLR statistic to reveal variation in the fit of ideal point models across individuals. If our goal is not to determine how many, but *which* Americans have preferences that are better described by ideal point models, then conditional averages of the LLR statistic are informative. While it would be a mistake to overinterpret any particular individual's LLR — because the distributions $F_d$ and $F_0$ will often overlap non-trivially, as seen in Figure 1 — this noise is not a factor once we can credibly estimate conditional expectations of the LLR. In particular, by regressing LLR values on covariates, we can identify which covariates are associated with higher or lower LLR values on average. We take this approach in section **??** to identify subsets of individuals that are better explained by ideal point models, such as politicians and sorted partisans.

# 4 Methods for Estimating and Evaluating Preference Models

We have argued that identifying whether or not an individual's preferences are better described by ideal points models than independent preferences is a matter of calculating log likelihoods under different implied probability distributions. The preceding discussion assumed fixed and known probability distributions, but in practice both independent preference and ideal point models have parameters that must be estimated from data. The objective of this section is to explain how we estimate these parameters and how we use cross-validation to avoid potential overfitting bias from estimating parameters and evaluating likelihoods on the same data sets.

## 4.1 Fitting Independent Preference Models

To estimate independent preference models, we do not need to concern ourselves with the relationships between preferences from different issue areas. Thus, estimating independent preference models is simply a matter of estimating an intercept parameter for each response option.

To make this explicit, suppose once again that the policy preferences of the $J$ issue areas are given by $x_i = (x_{i1}, \ldots, x_{iJ})$. The issue area response for issue $j$ might be categorical, so that $x_{ij} \in \{1, \ldots, K_j\}$. Because the issue areas are independent, this leads to $P_0(x_i = (k_1, \ldots, k_J)) = P_0(x_{i1} = k_1) \cdots P_0(x_{iJ} = k_J)$. Since each issue area is independent of the other issue areas, we only need to separately estimate the probability that $x_{ij} = k \in \{1, \ldots, K_j\}$ for each $j$, with the natural estimator being just the proportion of $j$ issue area preferences that equal $k$. This is a bit unstable, and interferes with the cross-validation estimator described in section 4.3, so we add half a pseudocount to each category. Adding half a pseudocount to each category is equivalent to putting Jeffrey's prior on each $P_0(x_{ij})$. This means that our final estimator of $P_0(x_{ij} = k)$ is

$$\hat{P}_0(x_{ij} = k) = \frac{0.5 + \sum_{i=1}^{N} I(x_{ij} = k)}{0.5K_j + N}. \tag{5}$$

### 4.2 Estimating Ideal Point Models

While we specify our ideal points in a manner consistent with the prior literature in political science, our estimation approach differs in two noticeable respects. First, as in the psychometrics literature, we cast the estimation as a marginal maximum likelihood problem, and, second, we estimate the models as if they are variational autoencoders, because this offers multiple improvements over the current state-of-the-art techniques in both political science and psychometrics, allowing us to better detect when ideal point models are suitable for any given individual.

Since our ideal point model specification is commonly used, we describe it only briefly. In the $d$-dimensional ideal point model, we suppose that given a latent variable $z_i \in \mathbb{R}^d$, known as the ideal point of individual $i$, the individual's stated policy preferences $(x_{i1}, \ldots, x_{iJ})$ are independent across issues. Thus,

$$P_d(x_i \mid z_i) = P_d(x_{i1} \mid z_i) \cdots P_d(x_{iJ} \mid z_i). \tag{6}$$

Since each issue depends on this common $z_i$ value, dependence between issues becomes apparent when we no longer condition on $z_i$. We put further structure on the problem by modeling each issue probability $x_{ij}$ given $z_i$ as a multinomial logistic regression. That is,

$$P_d(x_{ij} = k \mid z_i) = \frac{\exp\left(\alpha_{jk} + z_i'\beta_{jk}\right)}{\sum_{h=1}^{K_j} \exp\left(\alpha_{jh} + z_i'\beta_{jh}\right)}, \tag{7}$$

where $\theta = (\alpha, \beta)$ are the unknown parameters to be estimated.[8] This parametric model can be derived by assuming individual $i$ has quadratic utility over policies that is maximized at their ideal point $z_i$ with utility "errors" drawn from a type I extreme value distribution (this argument uses results from McFadden, 1973; Clinton, Jackman and Rivers, 2004).

[8] We ignore the problem of parameter identification. Parameters are only not identified if they have the same likelihood as other parameters, but since our only goal is to estimate likelihoods this does not present a problem for us.

In contrast with most of the political science literature, we cast the estimation of $\theta$ in terms of marginal maximum likelihood problem. Most applications of ideal point models in political science are concerned with estimating the latent ideal points $z_i$. However, we are not interested in estimating the ideal points $z_i$, just the likelihood $P_d(x_i)$.[9] We adopt the dominant framework used in the psychometrics literature for item response theory (IRT), and treat the estimation of $P_d$ as a marginal maximum likelihood (MML) problem. In particular, we assume that the ideal points are drawn from a $d$-dimensional multivariate standard normal distribution, $z_i \sim \mathcal{N}(0, I)$, and infer the implied marginal distribution of $x_i$ by integrating over our uncertainty in $z_i$.[10] The implied marginal distribution of $x_i$ is

$$P_d(x_i = (k_1, \ldots, k_J)) = \int P_d(x_i = (k_1, \ldots, k_J) \mid z_i) \, d\mathcal{N}(z_i \mid 0, I) \tag{8}$$

$$= \int \prod_{j=1}^{J} \left[ \frac{\exp\left(\alpha_{jk_j} + z_i' \beta_{jk_j}\right)}{\sum_{h=1}^{K_j} \exp\left(\alpha_{jh} + z_i' \beta_{jh}\right)} \right] d\mathcal{N}(z_i \mid 0, I). \tag{9}$$

Standard computational techniques in psychometrics for maximizing this likelihood with respect to $\theta$ involve approximating the integral with stochastic variants of the expectation-maximization (EM) algorithm (e.g., Chalmers, 2012).

However, we do not employ these stochastic EM methods to estimate $\theta$, and instead estimate the likelihood given in equations 8-9 with a (relatively shallow) variational autoencoder. Variational autoencoders (VAEs) are a generic method for fitting models with continuous latent variables (in our case, ideal points) in large data sets with intractable likelihoods or posterior distributions (Kingma and Welling, 2014). To briefly explain how we use them (see appendix A for more details), denote $p_\theta(x_i) = P_d(x_i)$. For VAEs, we introduce a variational distribution

[9] Jointly estimating the ideal points with joint maximum likelihood (JML) can lead to large-sample pathologies such as inconsistency, even for the $\theta$ parameters (Neyman and Scott, 1948; Ghosh, 1995). This is because each individual's ideal point can only be inferred from $J$ data points (i.e., the $J$ responses individual $i$ answers), and therefore inconsistent estimation of $z_i$ might prevent the consistent estimation of $\theta$.

[10] This corresponds to the prior distribution $z_i \sim \mathcal{N}(0, I)$ often employed in the empirical ideal point literature in political science. Note that the mean and covariance of the multivariate normal are not identified, so we can assume a zero mean and covariance matrix equal to the identity matrix without loss of generality. The only thing restrictive about the $z_i \sim \mathcal{N}(0, I)$ assumption is the shape (e.g., tail behavior) of the distribution.

$q_\phi(z_i \mid x_i)$, whose purpose is to approximate $p_\theta(z_i \mid x_i)$. It can be shown that

$$\begin{aligned}
&\log p_\theta(x_i) - D_{KL}(q_\phi(z_i \mid x_i) \,||\, p_\theta(z_i \mid x_i)) \\
&= E_{q_\phi(z_i|x_i)}[\log p_\theta(x_i \mid z_i)] - D_{KL}(q_\phi(z_i \mid x_i) \,||\, p_\theta(z_i)),
\end{aligned} \tag{10}$$

where $D_{KL}(q \,||\, p) \geq 0$ is the Kullback-Leibler (KL) divergence between $q$ and $p$, equal to $E_q[\log q - \log p]$, which in some sense measures how well $p$ approximates $q$. It can be shown that maximizing the right side of equation 10 with respect to $(\theta, \phi)$ is computationally quite easy. And by maximizing the right side of equation 10, we maximize the left side automatically. If we choose a rich enough class of distributions $q_\phi(z_i \mid x_i)$, then we can make $D_{KL}(q_\phi(z_i \mid x_i) \,||\, p_\theta(z_i \mid x_i)) \approx 0$, and thus maximizing the left side is approximately equivalent to maximizing the log likelihood $\log p_\theta(x_i)$.[11] This procedure gives us an approximate maximum likelihood estimator.[12]

We found that VAE-based estimators significantly outperformed the more traditional estimators found in political science and psychometrics. First, we document evidence in appendix B that VAEs achieve higher cross-validation log likelihoods, and thus have better model fit, on surveyed policy preferences. Since our ultimate goal requires calculating log likelihoods, this was enough for us to justify using VAEs. Second, also documented in appendix B, VAEs are far more successful at fitting multidimensional ideal point models, and so much so that they noticeably outperform unidimensional models in some surveys of the public. Finally, in our experience, VAEs tended to optimize faster than most stochastic EM algorithms for MML.[13]

[11] As discussed in appendix A, our $q_\phi(z_i \mid x_i)$ is a multivariate normal with a two-layer neural network as the mean and variance function of $x_i$.

[12] If we encounter missing data, we just assume that all data are missing at random (MAR) and that the parameters that determine missingness are different from $\theta$. Based on this, we only need to maximize the log likelihood of the observed data (see chapter 6 of Little and Rubin, 2014). For our purposes here, this assumption means we can ignore missing data terms in the sum $\log p_\theta(x_i \mid z_i)$.

[13] Although not faster than the exact EM and variational Bayes algorithms used in political science (Imai, Lo and Olmsted, 2016) for joint estimation. These methods, however, did not achieve competitive cross-validation likelihoods in comparison to the MML methods (see appendix B).

## 4.3 Avoiding Overfitting with Cross-validation

If we calculate log likelihoods on the same data we used to estimator our ideal point models, we might over-state the fit of our ideal point models, causing us to conclude that ideal point models fit more Americans than they do in reality.

To avoid this overfitting bias and get more realistic estimates $\hat{\Delta}(x_i)$ of the LLR, we use a 10-fold cross-validation scheme (Hastie, Tibshirani and Friedman, 2009, Ch. 7). In particular, before estimating any parameters $\theta$, we randomly split our data into 10 folds (subsets) of nearly equal size. For each individual $i$, this means we assign them a fold $v_i \in \{1, \ldots, V\}$ for $V = 10$. Then for each fold $v = 1, \ldots, V$ do the following steps:

1. Using data not in the $v$th fold, $\{x_i : v_i \neq v\}$, estimate $\hat{P}_0, \hat{P}_1, \ldots, \hat{P}_d$.
2. For each data point in the $v$th fold, $\{x_i : v_i = v\}$, and each dimension $d = 1, \ldots, D$, calculate $\hat{\Delta}_d(x_i) = \log \hat{P}_d(x_i) - \log \hat{P}_0(x_i)$.
3. For $s = 1, \ldots, S$ where $S$ is a large number (such as $10^3$), and for each $d = 1, \ldots, D$, draw $x^0_{vs} \sim \hat{P}_0$ and $x^d_{vs} \sim \hat{P}_d$. Use these values to compute $\hat{\Delta}_d(x^0_{vs})$ and $\hat{\Delta}_d(x^d_{vs})$.[14]

After doing this for each fold, we now have three sets of LLRs that we can use to estimate $\pi_d$, the proportion of Americans ideal point models are suitable for. The first set, $\{\hat{\Delta}_d(x^0_{vs}) : v = 1, \ldots, V, s = 1, \ldots, S\}$ are $V \times S = 10^4$ draws from the "null" distribution of LLRs when stated policy presences are independently drawn across issues. We can use this set to approximate $F_0$. Similarly, the second set $\{\hat{\Delta}_d(x^d_{vs}) : v = 1, \ldots, V, s = 1, \ldots, S\}$ can be used to approximate the $d$th "alternative" distribution $F_d$.[15] The final set $\{\hat{\Delta}_d(x_i) : i = 1, \ldots, N\}$ is the empirical

[14] When a data set has missing data, we need to make the $x^0_{vs}$, $x^d_{vs}$ values comparable to the real data $x_i$ by hiding some question responses. In particular, for each question $j$ we calculate the empirical missingness rate, and when we generate $x^0_{vs}$, $x^d_{vs}$ values we use this question missing rate to randomly determine whether the $j$th response of $x_{vs}$ is missing. Our strategy is inherently artificial and arbitrary (because the missing data mechanism is unknown), but when missingness is rare (as is typically the case in our data sources) we expect that our results are not affected too much. Note that the alternative mixture model estimates in appendix D, which are extremely similar to our main results, are robust to this artificial missing data generation, since those estimates do not require drawing any $x^0_{vs}$ or $x^d_{vs}$ values.

[15] We compute all estimates $\hat{F}$ based on a Gaussian kernel density estimator with the default `nrd0` bandwidth of `stats::density` in R.

distribution of LLRs. We use these distributions to estimate $\pi_d$ by finding how to weight the approximate null $\hat{F}_0$ and approximate $d$th alternative distribution $\hat{F}_d$ such that they equal the empirical distribution of LLRs, as described in section 3.2.

## 5 Data Sources

We use several data sets of policy preferences drawn from the public and politicians circa early 2010s. These data sets are based on typical uses in the public opinion and ideal point literatures, with policy preference questions developed by political scientists. We briefly summarize the data sources used below and list the specific policy preference questions used in the analysis in appendix C.

**2012 ANES.** We use questions from the 2012 American National Election Studies Time Series File, drawn from the replication material of Hill and Tausanovitch (2015). There are 5914 respondents and 29 policy questions. These questions cover a broad swath of politically salient topics, including health insurance, affirmative action, defense spending, immigration, welfare, and LGBT rights, as well as more generic questions about the role of government. About 9 percent of the response matrix is missing.

**2012 CCES.** We use "roll call" questions from the 2012 Cooperative Congressional Election Survey, where respondents are asked how they would vote on a series of bills that Congress also voted on. These data have been used to jointly scale Congress and the public (Bafumi and Herron, 2010), although we note that that is not our intention here. There are 52,619 respondents, answering 10 such questions on the 2012 CCES. The questions cover bills such as repealing the Affordable Care Act, ending Don't Ask Don't Tell, and authorization of the Keystone XL pipeline. Only about 2 percent of the response matrix is missing, mostly due to turnover in Congress.

We also match these survey responses to the corresponding roll-call votes in the House of Representatives and Senate. These votes took place in the 111th, 112th, and 113th Congresses.

We match 8 questions to roll-call votes.[16] A full list of the votes used for ideal point estimation is available in appendix C. About 37 percept on the roll call matrix is missing.

**Broockman (2016) Survey.** Finally, we use data from Broockman's (2016) survey of sitting state legislators. This survey contains responses from 225 state legislators on 31 policy questions. Question topics include Medicare, immigration, gun control, tax policy, gay marriage, and medical marijuana, among others. Only about 5 percent of the response matrix is missing. Additionally, we use a paired survey of the public that is also reported in Broockman (2016). The 31 questions in this data are identical to those asked of state legislators. We only use the first wave of the survey, in which there are 997 respondents and no missing data.

# 6 Results

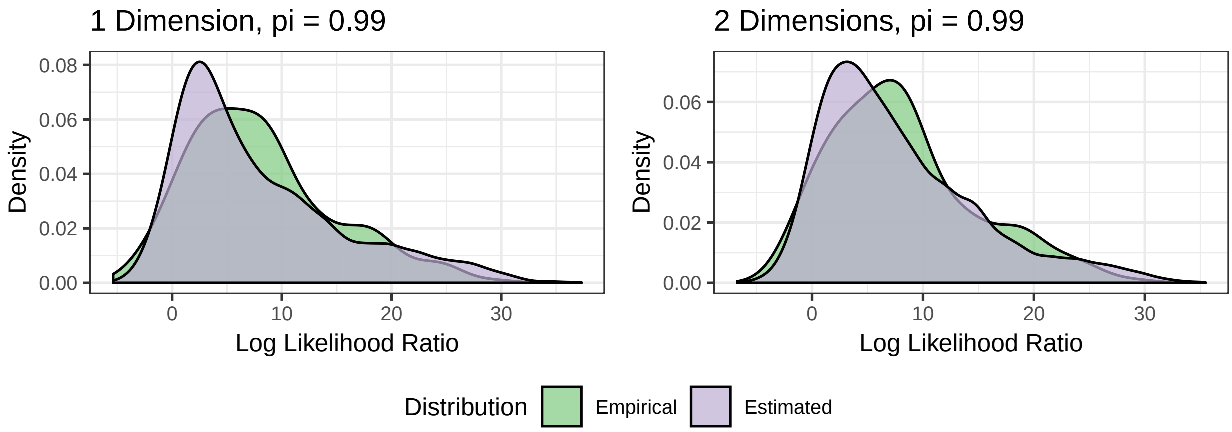
## 6.1 How Many Americans Hold Related Political Attitudes?

We are interested first in the proportion of the population that holds attitudes that are better summarized by an ideal point model that allows preferences to be related than an independent-preference model that treats attitudes as independent. Our approach, as previously outlined, is to treat each member of the population as coming from either the null (independent-response) model, or an alternative $d$-dimensional model, then estimating a mixing parameter $\pi_d$ that describes the proportion of "ideal point" types.

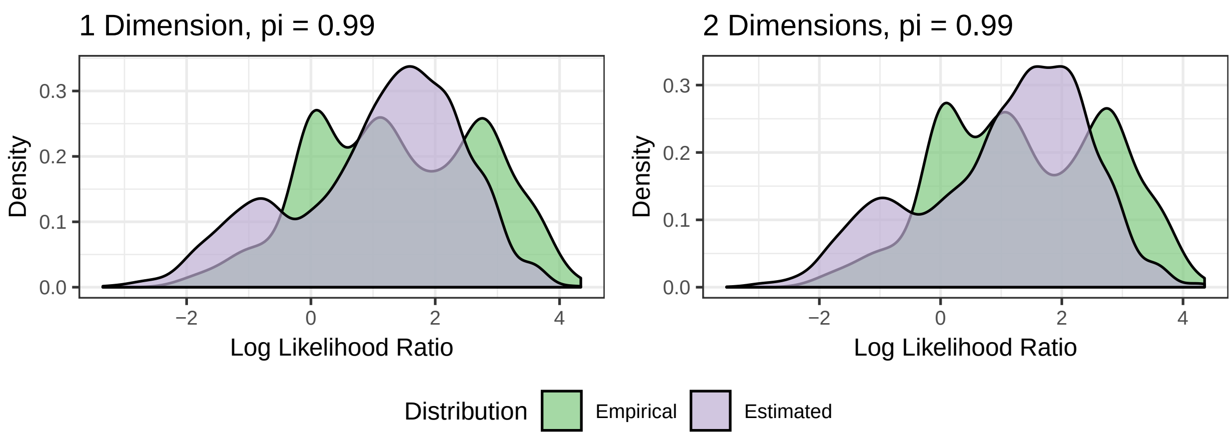
Before we present our main results that focus on the public, we first present results of the procedure applied to data from politicians. We have strong a priori reason to think that politicians' preferences are highly organized along ideological lines. As such, we expect to estimate that nearly all politicians are better characterized by ideal point models, implying that we should estimate $\pi_d \approx 1$.

Indeed, this is exactly what we find. Figure 2 plots the distribution of the individual-level

[16] The excluded "roll calls" are CC332B, the Simpson-Bowles Budget Plan, and CC332E, Birth Control Exemption, for which we were unable to find floor votes in both chambers.




(a) State Legislators: Broockman (2016) Survey




(b) Roll Call Votes: Paired 2012 CCES Questions

Figure 2: Holdout log likelihood ratio (LLR) associated with 1-dimensional (left) and 2-dimensional (right) models, relative to an unrelated-preference model. The input data are: (panel a) 31 survey questions answered by sitting state legislators, collected by Broockman (2016); (panel b) 8 roll call votes in the House and Senate that correspond to CCES roll-call questions. The green density estimate shows the empirical LLR distribution, while the purple density estimate shows the estimated LLR distribution under the best-fitting mixture model. Imperfect fit in the CCES roll call votes could be attributable to either the low number of questions (eight) or the relatively high degree of missing data (roughly 37 percent of votes).

LLR statistic associated with 1- and 2-dimensional ideal point models, relative to an unrelated-preference model, using state legislator survey data from Broockman (2016) and Congressional votes that correspond to CCES roll call survey questions. In all cases, we estimate $\hat{\pi} \approx 0.99$, indicating that for virtually all politicians, an ideal point model is a better representation of their responses than the unrelated-preference model. There is no debate that politicians have highly

| Population | Data Source | $\hat{\pi}_1$ | $\hat{\pi}_2$ |
|---|---|---|---|
| *Politicians:* | State Legislator Survey | 0.99 | 0.99 |
| | | (0.012) | (0.036) |
| | Roll Call Votes | 0.99 | 0.99 |
| | | (0.034) | (0.034) |
| *Public:* | ANES 2012 | 0.93 | 0.98 |
| | | (0.006) | (0.005) |
| | CCES 2012 Roll Calls | 0.92 | 0.97 |
| | | (0.004) | (0.003) |
| | Broockman (2016) Survey | 0.74 | 0.87 |
| | | (0.027) | (0.022) |

Table 2: Estimates of $\pi_1$ and $\pi_2$ across data sources, with standard errors shown in parentheses. $\pi_d$ gives the proportion of the population that is better described by a $d$-dimensional ideal point model compared to an unrelated-preference model.

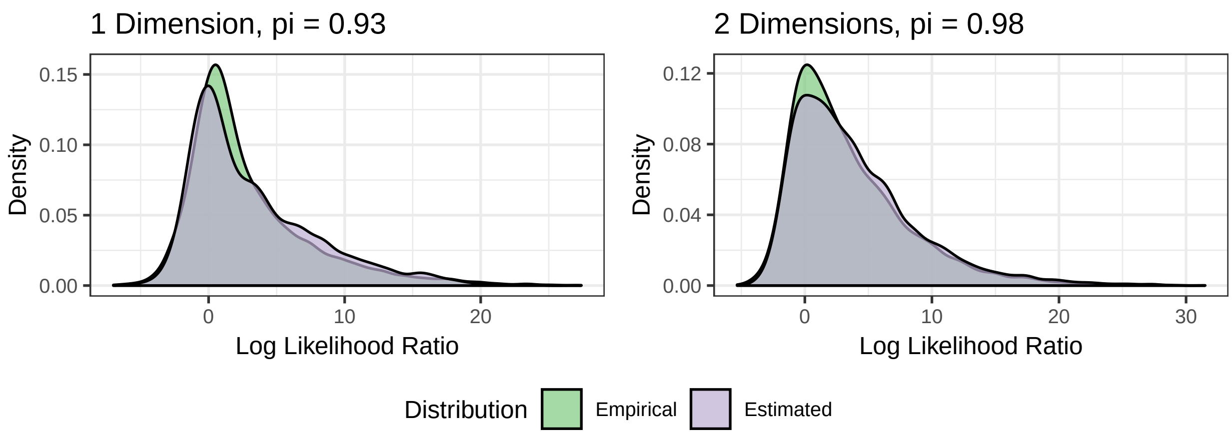
structured attitudes, so this result bolsters our confidence in our mixture-model approach.

We now turn to our full set of $\pi$ estimates, across both politicians and the public. These estimates for 1- and 2-dimensional models, along with associated standard errors, are reported in Table 2.[17] The first two rows repeat the results for politicians discussed above. The rest of the table suggests that the vast majority of the public has political attitudes that are better described by an ideal point model than an unrelated-preference model.

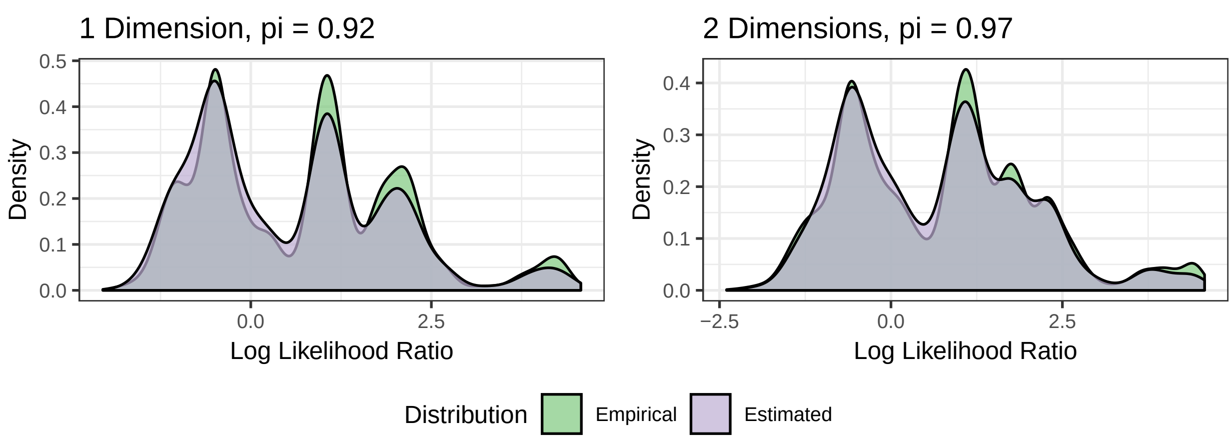
Consider the results for the 2012 ANES. Recall that there are 29 policy questions in this dataset, covering a broad range of politically relevant issues. We estimate that about 93% of ANES respondents give answers more consistent with a 1-dimensional ideal point model than an independent-preference model. This estimate increases to 98% if we allow the ideal point model to have 2-dimensions, consistent with Ansolabehere, Rodden and Snyder (2008) and Treier and Hillygus (2009).

We obtain similar estimates when we apply our method to the 10 "roll call" questions asked on the CCES in 2012. Using this sample, we estimate that 92 and 97% of respondents are better characterized by a 1- and 2-dimensional ideal point model, respectively, relative to the

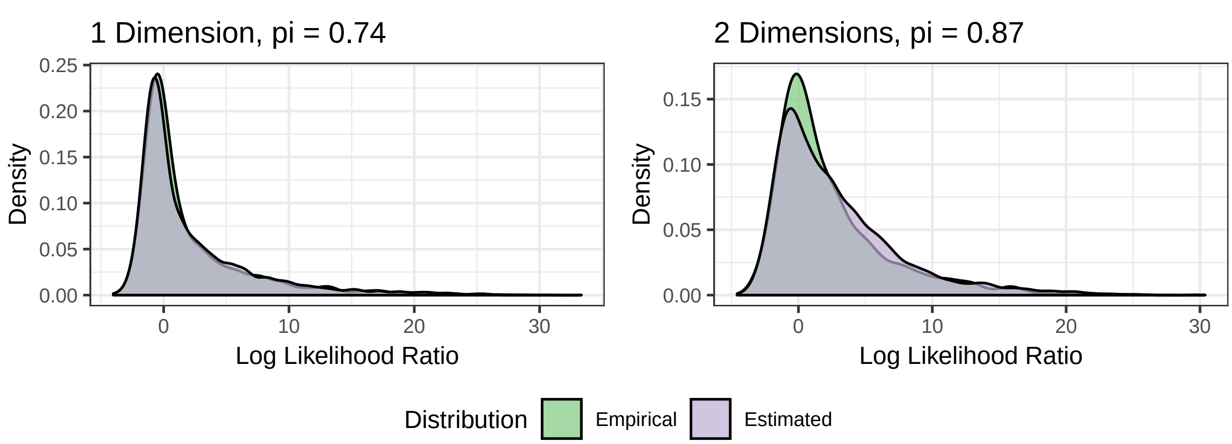
[17] Standard errors are estimated using the inverse Fisher information and are conditional on the LLR density estimates.




(a) ANES 2012




(b) CCES 2012 Roll Calls




(c) Broockman (2016) Survey

Figure 3: Holdout log likelihood ratio (LLR) associated with 1-dimensional (left) and 2-dimensional (right) models, relative to an unrelated-preference model. The input data are: (panel a) the 2012 ANES; (panel b) the CCES roll call items from 2012; (panel c) survey responses from the public collected by Broockman (2016). The green density estimate shows the empirical LLR distribution, while the purple density estimate shows the estimated LLR distribution under the best-fitting mixture model.

unrelated-preference model. This finding is notable for two reasons. First, there are many fewer questions here than in the ANES, suggesting that it is not necessary to ask a large number of questions to detect structure in survey responses. Second, these questions were chosen to mirror actual roll call votes in Congress. While citizens may not be well-informed about these issues (cf. Hill and Huber 2017), it is still possible to extract meaningful information from the joint distribution of responses.

The outlier in our results comes from the survey data collected by Broockman (2016). In this dataset, we estimate that about three-quarters of the population is better described by a 1-dimensional model. When we allow for a second dimension, the estimate increases to 87%. While these estimates are substantially lower than for the other data sources, they still indicate that a large majority of citizens hold meaningful preferences.

One might reasonably ask whether our mixture-model approach provides a good approximation of the empirical distribution of log likelihood ratios of $d$-dimensional models relative to unrelated-preference models. If it turns out that the empirical LLR distribution looks completely dissimilar to the estimated mixture distribution, then we might view the $\hat{\pi}_d$ estimates with skepticism. Figure 3 therefore plots the empirical and estimated LLR distributions for each mass public data source. There is a very good fit between the estimated mixture distributions and the empirical distributions. Even for the CCES data, which contains only 10 survey questions and therefore has a very lumpy LLR distribution, the estimated distribution mirrors closely the empirical distribution. These visual results provide further support of the validity of our procedure.

Overall, these results strongly support the hypothesis most citizens hold political attitudes that are related across issues, just as politicians do. However, it is important to note that the increase in fit associated with ideal point models is substantially higher for elite samples than for the public. Direct evidence of this can be seen in Figure 4. This plot shows the distribution of average changes in the probability of observing an individual question response under the ideal point models compared to an unrelated-preference model, using data from both

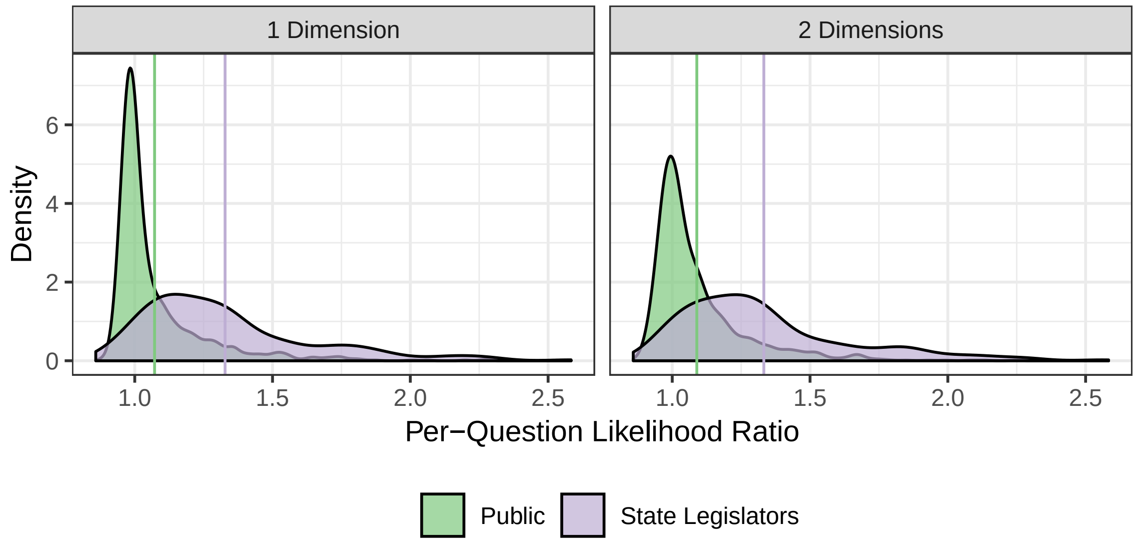



Figure 4: Distribution of per-question likelihood ratio under ideal point models, relative to unrelated preference models, for both the public and politician samples from the Broockman (2016) survey. Each observation is an individual respondent. The $x$-axis statistic is calculated as $\exp(LLR/31)$, where we divide by 31 because there are 31 questions in the survey. Vertical lines show the means for each sample.

the public and elite samples from Broockman (2016).[18] For example, a value of 2 indicates that for an individual respondent, we are twice as likely to observe their actual response to any given question, on average, under the alternative model than under the null model. These two samples — public and politicians — answered the same set of questions, so the holdout log-likelihood ratios are directly comparable.

The distribution The increase in fit for ideal point models is substantially higher among the politician sample than the mass public. The average politician response is 1.33 times more likely under a 1-dimensional model than an unrelated-preference model. In contrast, the average response from the mass public is 1.07 times more likely under a 1-dimensional model than under an unrelated-preference model. Thus, while the 1-dimensional model better characterizes both sets of actors, the improvement in fit is substantially larger for politicians than for the public.

[18]We calculate this statistic as $\exp(LLR/31)$, where we divide by 31 because there are 31 questions in the sample.

## 6.2 What Explains Related Preferences?

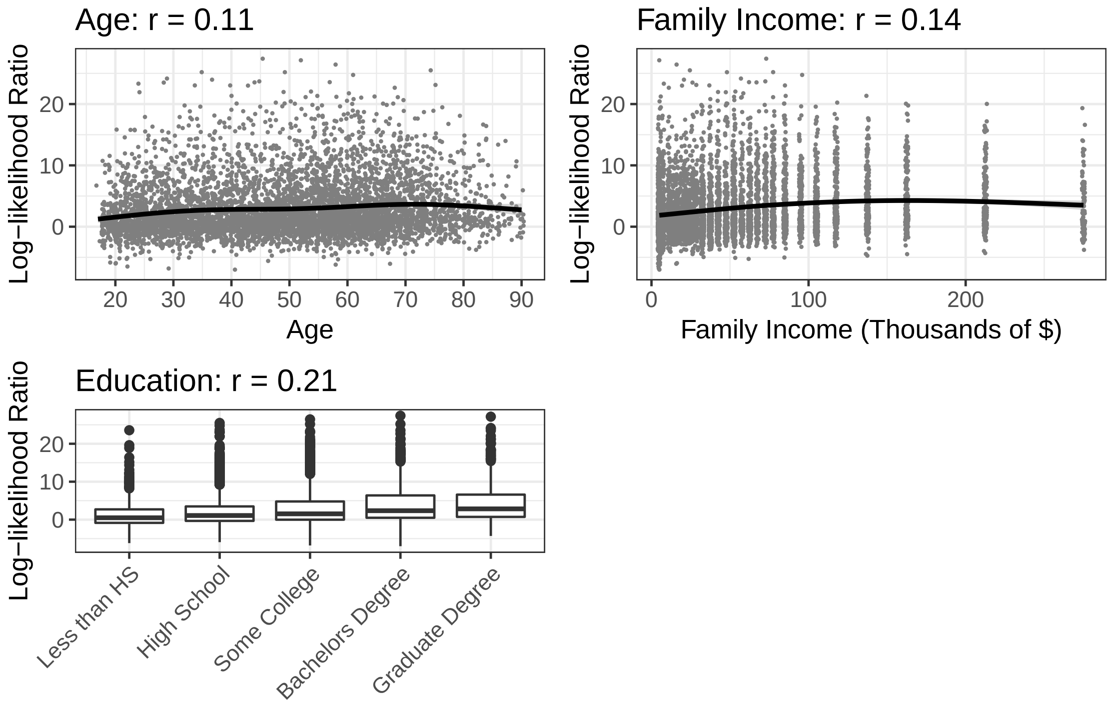
Our results suggest that most people are better described as holding political attitudes that are related across issues than as holding independent attitudes. But there is considerable variation across members of the public in just how much better the ideal point model fits than the unrelated preference model. The question of exactly *who* has "meaningful" preferences is the subject of a large literature. Recent work has pointed to political knowledge as a key determinant of having meaningful preferences (Barber and Pope, 2017; Freeder, Lenz and Turney, 2018). Other work suggests that strong partisans (Jacobson, 2012), the politically engaged (Abramowitz, 2012), the highly educated, or people who are "sorted" partisans (Levendusky, 2009; Sniderman and Stiglitz, 2012) are likely to have more meaningful preferences.

In this section, we use our measure of preference structure to tackle this question. We find that it is not just one variable that predicts how related an individuals' attitudes are. Instead, we find that people with high levels of political knowledge, sorted partisans, political donors, and Republicans all have preferences that tend to go together. Sociodemographic variables, such as age, income, and education, are not as strongly predictive of having related preferences.

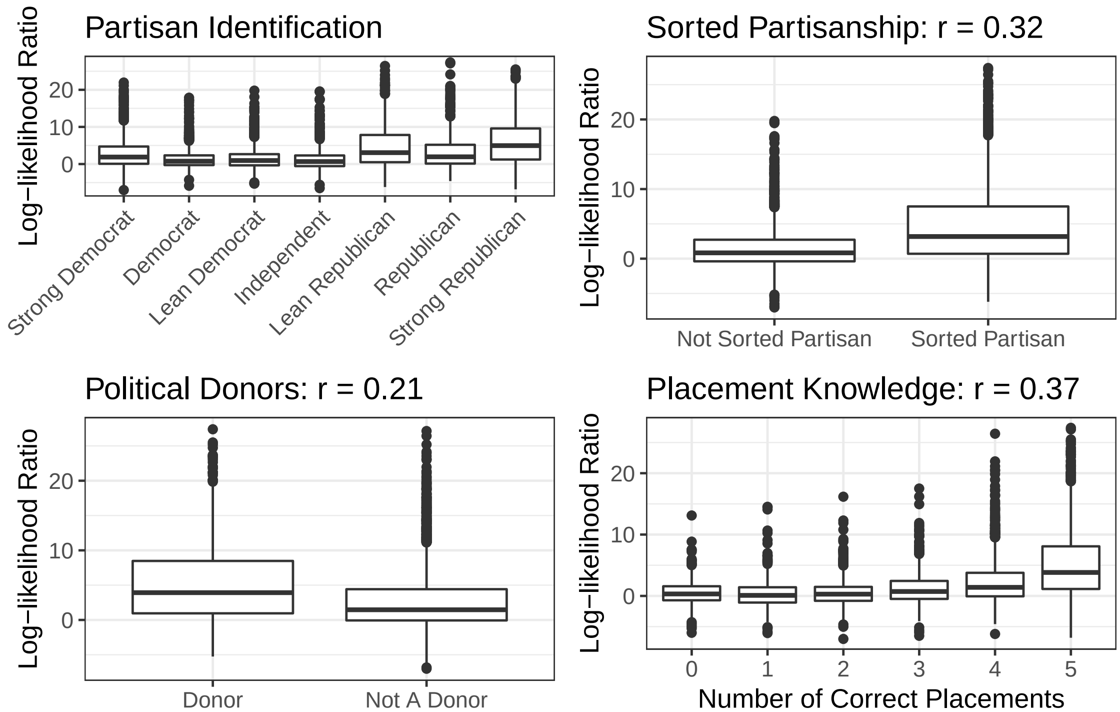
Our empirical strategy in this section is to regress individual-level log-likelihood ratios under a 1- or 2-dimensional model on sociodemographic and political covariates drawn from the ANES. Positive coefficients indicate that a higher value of the covariate is associated with a larger log-likelihood ratio, indicating that the ideal point model gives a larger increase in fit for that individuals' responses over the unrelated preference model.

Coding the sociodemographic variables, which include age, family income, and education, is straightforward.[19] We construct several political variables as follows. First, we take the standard 7-point partisan identification scale. We code respondents as "sorted partisans" if they self-identify as liberal (conservative) on a 7-point ideological self-placement scale *and* they self-identify as being a Democrat (Republican). We include "leaners" as self-identified partisans. If respondents place themselves at the middle of either scale, they are coded as not

[19] Family income is given in relatively narrow bins; we recode this variable as the midpoint of the bin.




(a) Sociodemographic Variables




(b) Political Variables

Figure 5: These figures show the bivariate correlations between the holdout log-likelihood ratio of a 1-dimensional ideal point model and (panel A) sociodemographic variables and (panel B) political variables. Correlations are linear correlation for continuous and binary variables (age, income, political donor, sorted partisans) and rank-order correlations for ordered variables (education, placement knowledge).

being sorted partisans. Next, we include an indicator for having donated to a candidate, a party, or another political organization in the contemporary election cycle. We take this variable as a proxy for being highly engaged in politics.

Finally, we follow the approach of Freeder, Lenz and Turney (2018) in constructing a political knowledge variable. The ANES asks respondents to order the presidential candidates on a left-right scale on five issues.[20] We count a placement as being correct if respondents indicate that Barack Obama is more liberal than Mitt Romney. As our measure of knowledge, we simply count the number of correct placements.

To begin, in Figure 5 we plot the bivariate relationship between the individual-level holdout LLR and each covariate. The top panel shows the sociodemographic covariates, and the bottom panel shows the political covariates. For categorical covariates, we use boxplots to visualize the distribution of LLR; for continuous variables we show scatter plots.

First, we note that there is a large amount of variation across all levels of the covariates. In short, there is no single group that uniformly high or low LLR — again emphasizing the importance of studying continuous, individual-level measures of related preferences rather than qualitative and aggregate measures. Second, most variables are at least mildly correlated with average LLR. There are increases in average LLR for sorted partisans, compared to unsorted partisans, for political donors, and for people who exhibit a high degree of political knowledge. There is also generally a $U$-shaped relationship between partisan identification and LLR, suggesting that people who self-identify as partisans are better described by ideal point models than people who self-identify as independents. In contrast, the relationships between the sociodemographic variables and the LLR measure are generally weak. Age and family income only (linearly) account for about 1-2% of the variation in LLR, while education accounts for only about 4%.

These bivariate results are only somewhat informative, given that many of these variables are correlated with each other. To study the correlations with more nuance, we regress the LLR

[20] These issues are: a government spending/services scale; a defense spending scale; a government health insurance scale; a guaranteed jobs scale; and an aid-to-Blacks scale.

| | | 1D LLR | 2D LLR |
|---|---|---|---|
| *Sociodemographic variables:* | Age | $0.044^{*}$ | $0.058^{**}$ |
| | | (0.020) | (0.021) |
| | Age squared | −0.0004 | $-0.0005^{*}$ |
| | | (0.0002) | (0.0002) |
| | Family income | 0.002 | 0.002 |
| | | (0.003) | (0.003) |
| | Family income squared | −0.00001 | −0.00001 |
| | | (0.00001) | (0.00001) |
| | Educ: High school | 0.187 | −0.205 |
| | | (0.207) | (0.226) |
| | Educ: Some college | $0.451^{*}$ | −0.042 |
| | | (0.206) | (0.223) |
| | Educ: Bachelors degree | $0.629^{**}$ | 0.323 |
| | | (0.239) | (0.255) |
| | Educ: Graduate degree | $0.554^{*}$ | 0.398 |
| | | (0.275) | (0.291) |
| *Political variables:* | Party: Dem. | $-0.903^{***}$ | $-1.058^{***}$ |
| | | (0.162) | (0.177) |
| | Party: Lean Dem. | $-0.769^{***}$ | $-0.923^{***}$ |
| | | (0.174) | (0.184) |
| | Party: Independent | 0.252 | −0.120 |
| | | (0.188) | (0.200) |
| | Party: Lean Rep. | $1.799^{***}$ | $1.051^{***}$ |
| | | (0.264) | (0.264) |
| | Party: Rep. | $0.659^{**}$ | −0.070 |
| | | (0.224) | (0.226) |
| | Party: Strong Rep. | $2.443^{***}$ | $1.597^{***}$ |
| | | (0.240) | (0.243) |
| | Sorted partisan | $1.515^{***}$ | $1.491^{***}$ |
| | | (0.125) | (0.130) |
| | Political donor | $1.540^{***}$ | $1.596^{***}$ |
| | | (0.207) | (0.211) |
| | Num. correct placements | $0.814^{***}$ | $0.762^{***}$ |
| | | (0.039) | (0.041) |
| | Constant | $-2.979^{***}$ | $-1.879^{***}$ |
| | | (0.512) | (0.548) |
| | $N$ | 4,825 | 4,825 |
| | $R^2$ | 0.254 | 0.213 |

Table 3: Regression of sociodemographic and political variables on the holdout log-likelihood ratio of a 1-dimensional and 2-dimensional ideal point models compared to an unrelated-preference model, using data from the 2012 ANES. Higher values indicate that the ideal point model better describes an individuals' preferences. Income is measured is thousands of dollars; the omitted category for education is less than a high school diploma; the omitted category for partisanship is strong Democrat; "sorted partisan" takes the value of 1 if the respondent is liberal Democrat or conservative Republican and 0 otherwise; "number of correct placements" refers to the number of issues on which the respondent could correctly order the 2012 presidential candidates and ranges from 0 to 5. Heteroskedasticity-robust standard errors are shown in parentheses. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

on the full set of covariates. The results are shown in Table 3. The first column uses the LLR for the 1-dimensional model as the left-hand-side variable; the second column uses the LLR for the 2-dimensional model.

The results generally confirm the suspicion that the political variables are highly correlated with increase in fit for an ideal point model. Starting at the top of the table, we see that age and income are only weakly related to the LLR. Education appears to be mildly positively related to the LLR in the 1-dimensional case, but the magnitude of the coefficients decreases and they become insignificant when we consider the LLR for the 2-dimensional case.

Turning now to the partisan identification coefficients, an interesting pattern emerges: across the board, Republicans' attitudes tend to "go together" more than Democrats' attitudes. The omitted category in the regression table is strong Democrat. Nearly every category of Republicans (save for weak Republicans in the 2-dimensional case) has a positive and statistically significant coefficient. Thus, even after accounting for political knowledge and engagement, as well as demographic characteristics, Republicans are better described as ideological than Democrats. This finding fits with the argument, advanced by Grossmann and Hopkins (2016), that that Republicans are bound together by ideology while Democrats are bound together by interest groups.

Next, less surprisingly, respondents who are sorted partisans are better described by ideal point models than unsorted partisans — a finding that fits in with the observation dating back to Converse (1964) that people who understand politics in the same way as political elites are more likely to hold meaningful preferences.

Finally, we find that there is a strong relationship between political knowledge, measured as the number of correct candidate placements, and the fit of ideal point models. This finding is line with recent work by Sniderman and Stiglitz (2012), Freeder, Lenz and Turney (2018), and Barber and Pope (2017), who also find that political knowledge is a strong predictor of over-time attitude stability and cross-sectional correlations among attitudes.

Overall, this multivariate analysis reveals that there is no single variable that overwhelms the others in predicting how closely an individuals' attitudes "go together." Indeed, there is substantial heterogeneity in the population, and several political variables are independently correlated with having preferences that are best described by an ideal point model.

Finally, a cautionary note. While our approach allows us to characterize which members of the public are better described by an ideal point model, it does not allow us to identify *why* the ideal point model fits better. The most straightforward explanation is that some members of the public simply have preferences that are more structured; our regression results, which show that high-knowledge individuals are more constrained, suggest that this is part of the story. However, another source of variation is heterogeneity in measurement error. If surveys are a worse tool at measuring attitudes for some people than for others, it would have the effect of depressing the LLR for those people. Our method does not allow us to disentangle these two sources of heterogeneity.

# 7 Conclusion

How many and which Americans hold policy preferences that are related across issues? While this may seem like a simple question, a focus on aggregate measures, such as correlations, has prevented the field of public opinion from reaching a firm conclusion. On one hand, low interitem correlations shows that stated preferences among the public is low (Converse, 1964), but this could be due both to few people caring about politics, measurement error in survey responses (Achen, 1975), or more complex relationships between attitudes that cannot be detected with bivariate linear relationships. On the other hand, combining policy preferences into ideal points or similar ideology measures may remove measurement error and help predict voting behavior (Ansolabehere, Rodden and Snyder, 2008), but that does not tell us whether measurement error applies uniformly to all Americans or the proportion of Americans who can be justifiably summarized by ideal point measures.

To identify how many and for which Americans the measurement error story is most plausible, we develop an individual-level measure that responds positively if ideal point models explain an individualâĂŹs survey response better than a model that assumes preferences are unrelated. Using this new measure and a mixture model over ideal point types and independent issue types, we find that the measurement error account of survey response applies to the vast majority of Americans. Over 90% of Americans express attitudes that are better explained by an ideal point model than an independent-preference model. However, that is not to say that there is no variation in ideal point model fit. To the contrary, we find that ideal point models are much better predictions of politicians' stated attitudes than non-politicians. We also find that, among the public, there are strong correlations between ideal point model fit and several political variables, including political knowledge, political engagement, sorted partisanship, and being a Republican. In contrast, sociodemographic variables are relatively weak predictors of ideal point model fit.

On the whole, our findings suggest that ideal point models, and in general measures that use the joint distribution of survey responses to infer attitudes, are justified for most (although not all) Americans. However, this does not mean that ideal points are the only measure of attitudes worth observing. Recent work has demonstrated that relying solely on estimated ideal points can mask important individual-level idiosyncrasies that are stable over time (Broockman, 2016; Lauderdale, Hanretty and Vivyan, 2017). A challenge for future work is to extract individual preference measures over issues that simultaneously guard against measurement error, like ideal points, without masking real (stable) individual-level deviations from common ideological structures.

Another natural extension of this work is disentangling the two main sources of variation in ideal point fit: measurement error and individual heterogeneity. Our regression results show how sociodemographic and political observables predict ideal point fit, but they do not tell us how much of the variance explained by these variables is variance arising from measurement error or variance from natural heterogeneity in the population of citizens over how well their attitudes fit together. The latter variation is more interesting politically; however, with cross-sectional data it does not seem possible to disentangle these two sources of variation.

# References

Abramowitz, Alan. 2012. *The Disappearing Center: Engaged Citizens, Polarization, and American Democracy*. New Haven, Connecticut: Yale University Press.

Achen, Christopher H. 1975. "Mass Political Attitudes and the Survey Response." *American Political Science Review* 69(4):1218–1231.

Ansolabehere, Stephen, Jonathan Rodden and James M Snyder. 2006. "Purple America." *Journal of Economic Perspectives* 20(2):97–118.

Ansolabehere, Stephen, Jonathan Rodden and James M. Snyder. 2008. "The Strength of Issues: Using Multiple Measures to Gauge Preference Stability, Ideological Constraint, and Issue Voting." *American Political Science Review* 102(02):215–232.

Bafumi, Joseph and Michael C. Herron. 2010. "Leapfrog Representation and Extremism: A Study of American Voters and Their Members in Congress." *American Political Science Review* 104(03):519–542.

Barber, Michael J. and Jeremy C. Pope. 2017. "Who is Constrained? Measuring Issue Consistency in the American Public.".

Benjamini, Yoav and Yosef Hochberg. 1995. "Controlling the false discovery rate: a practical and powerful approach to multiple testing." *Journal of the Royal Statistical Society* 57(1):289–300.

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

Chalmers, Philip R. 2012. "mirt: A Multidimensional Item Response Theory Package for the R Environment." *Journal of Statistical Software* 48(6):1–29.

Clinton, Joshua D, Simon Jackman and Douglas Rivers. 2004. "The Statistical Analysis of Roll Call Data." *American Political Science Review* 98(2):355–370.

Converse, Philip E. 1964. The Nature of Belief Systems in Mass Publics. In *Ideology and Discontent*, ed. David Apter. New York: The Free Press pp. 206–261.

Delli Carpini, Michael X. and Scott Keeter. 1997. *What Americans Know About Politics and*

*Why It Matters*. New Haven: Yale University Press.

Efron, Bradley. 2004. "Large-scale simultaneous hypothesis testing: The choice of a null hypothesis." *Journal of the American Statistical Association* 99(465):96–104.

Freeder, Sean, Gabriel S. Lenz and Shad Turney. 2018. "The Importance of Knowing "What Goes with What": Reinterpreting the Evidence on Policy Attitude Stability." *Journal of Politics* 81(1).

Ghosh, Malay. 1995. "Inconsistent maximum likelihood estimators for the Rasch model." *Statistics and Probability Letters* 23(2):165–170.

Goplerud, Max. 2018. "A Multinomial Framework for Ideal Point Estimation." *Political Analysis* .

Grossmann, Matt and David A. Hopkins. 2016. *Asymmetric Politics: Ideological Republicans and Group Interest Democrats*. New York: Oxford University Press.

Hastie, Trevor, Robert Tibshirani and Jerome Friedman. 2009. *Elements of Statistical Learning: Data Mining, Inference, and Prediction*. 2 ed. New York: Springer.

Hill, Jennifer L. and Hanspeter Kriesi. 2001. "An Extension and Test of Converse's "Black-and-White" Model of Response Stability." *American Political Science Review* 95(2):397–413.

Hill, Seth J and Chris Tausanovitch. 2015. "A Disconnect in Representation? Comparison of Trends in Congressional and Public Polarization." *Journal of Politics* 77(4):1058–1075.

Hill, Seth J. and Gregory A. Huber. 2017. "On The Meaning of Survey Reports of Roll Call Votes Not Cast in a Legislature." *American Journal of Political Science* (Forthcoming).

Imai, Kosuke, James Lo and Jonathan Olmsted. 2016. "Fast Estimation of Ideal Points with Massive Data." *American Political Science Review* 110(4):1–20.

Jacobson, Gary C. 2012. "The Electoral Origins of Polarized Politics: Evidence From the 2010 Cooperative Congressional Election Study." *American Behavioral Scientist* 56(12):1612–1630.

Jessee, Stephen A. 2009. "Spatial Voting in the 2004 Presidential Election." *American Political*

*Science Review* 103(01):59.

Kinder, Donald R. 2003. Belief Systems After Converse. In *Electoral Democracy*, ed. Michael MacKuen and George Rabinowitz. Ann Arbor: University of Michigan Press.

Kingma, Diederik P and Max Welling. 2014. Auto-Encoding Variational Bayes. In *International Conference on Learning Representations (ICLR)*.

Lauderdale, Benjamin E., Chris Hanretty and Nick Vivyan. 2017. "Decomposing Public Opinion Variation into Ideology, Idiosyncrasy and Instability." *Journal of Politics* pp. 1–9.

Lehmann, Erich L and Joseph P Romano. 2006. *Testing Statistical Hypotheses*. Springer.

Levendusky, Matthew S. 2009. *The Partisan Sort: How Liberals Became Democrats and Conservatives Became Republicans*. Chicago: University of Chicago Press.

Little, Roderick J A and Donald B Rubin. 2014. *Statistical Analysis with Missing Data*. 2nd ed. Hoboken, NJ: John Wiley & Sons.

McFadden, Daniel. 1973. "Conditional logit analysis of qualitative choice behavior.".

Neyman, J and Elizabeth L Scott. 1948. "Consistent Estimates Based on Partially Consistent Observations." *Econometrica* 16(1):1–32.

Pan, Jennifer and Yiqing Xu. 2017. "China's Ideological Spectrum." *Journal of Politics* .

Sniderman, Paul M. and Edward H. Stiglitz. 2012. *The Reputational Premium: A Theory of Party Identification and Policy Reasoning*. Princeton: Princeton University Press.

Sullivan, John L., James E. Piereson and George E. Marcus. 1978. "Ideological Constraint in the Mass Public: A Methodological Critique and Some New Findings." *American Journal of Political Science* 22(2):233–249.

Tausanovitch, Chris and Christopher Warshaw. 2013. "Measuring Constituent Policy Preferences in Congress, State Legislatures, and Cities." *Journal of Politics* 75(02):330–342.

Treier, S. and D. S. Hillygus. 2009. "The Nature of Political Ideology in the Contemporary Electorate." *Public Opinion Quarterly* 73(4):679–703.

Zaller, John and Stanley Feldman. 1992. "A Simple Theory of the Survey Response : Answering Questions versus Revealing Preferences." *American Journal of Political Science*

36(3):579–616.

# Appendix

## Table of Contents

# A Ideal Point Models as Shallow Variational Autoencoders

# B Comparing Ideal Point Estimators

# C Data Appendix

### Broockman (2016) State Legislator Variables

| Variable | Description |
|---|---|
| iq_vouchers | The government should provide parents with vouchers to send their children to any school they choose, be it private, public, or religious. (Binary) |
| iq_medicalpot | Allow doctors to prescribe marijuana to patients. (Binary) |
| iq_taxesover250k | Increase taxes for those making over $250,000 per year. (Binary) |
| iq_overturnroe | Overturn Roe v. Wade. (Binary) |
| iq_privitsocialsec | Allow workers to invest a portion of their payroll tax in private accounts that they can manage themselves. (Binary) |
| iq_gaymarriage | Same-sex couples should be allowed to marry. (Binary) |
| iq_unihealth | Implement auniversal health care program to guarantee coverage to all Americans, regardless of income. (Binary) |
| iq_medlawsuits | Limit the amount of punitive damages that can be awarded in medical malpractice lawsuits. (Binary) |
| iq_guncontrol | There should be strong restrictions on the purchase and possession of guns. (Binary) |
| iq_illegalim | Illegal immigrants should not be allowed to enroll in government food stamp programs. (Binary) |
| iq_enda | Include sexual orientation in federal anti-discrimination laws. (Binary) |
| iq_affaction | Prohibit the use of affirmative action by state colleges and universities. (Binary) |
| iq_unfunding | The US should contribute more funding and troops to UN peacekeeping missions. (Binary) |
| iq_fundarts | The government should not provide any funding to the arts. (Binary) |
| iq_dealthpenalty | I support the death penalty in my state. (Binary) |
| iq_repealcapgainstax | Repeal taxes on interest, dividends, and capital gains. (Binary) |
| iq_epaprohibit | Prohibit the EPA from regulating greenhouse gas emissions. (Binary) |
| iq_birthcontrolmandate | Health insurance plans should be required to fully cover the cost of birth control. (Binary) |

| | |
|---|---|
| iq_subsidizeloans | The federal government should subsidize student loans for low income students. (Binary) |
| eq_guns | Which statement comes closest to describing your views on gun control? (1-7 scale) |
| eq_health | Which statement comes closest to describing your views on the issue of health care? (1-7 scale) |
| eq_immigration | Which statement comes closest to describing your views on immigration? (1-7 scale) |
| eq_taxes | Which statement comes closest to describing your views on taxes? (1-7 scale) |
| eq_abortion | Which statement comes closest to describing your views on abortion? (1-7 scale) |
| eq_environment | Which statement comes closest to describing your views on pollution and the environment? (1-7 scale) |
| eq_medicare | Which statement comes closest to describing your views on Medicare, the government's program for covering the elderly's health care costs? (1-7 scale) |
| eq_gays | Which statement comes closest to describing your views on rights for gays and lesbians? (1-7 scale) |
| eq_affirmativeaction | Which statement comes closest to describing your views on affirmative action in higher education? (1-7 scale) |
| eq_unions | Which statement comes closest to describing your views on unions? (1-7 scale) |
| eq_education | Which statement comes closest to describing your views on public funding for private school education? (1-7 scale) |
| eq_contraception_version2 | Which statement comes closest to describing your views on birth control? (1-7 scale) |

## 2012 ANES Variables, from Hill and Tausanovitch (2015)

| Variable | Description |
|---|---|
| VCF0806 | R Placement: Government Health Insurance Scale |
| VCF0809 | R Placement: Guaranteed Jobs and Income Scale |
| VCF0823 | R Opinion: Better off if U.S. Unconcerned with Rest of World |
| VCF0830 | R Placement: Aid to Blacks Scale |
| VCF0838 | R Opinion: By Law, When Should Abortion Be Allowed |
| VCF0839 | R Placement: Government Services/Spending Scale |
| VCF0843 | R Placement: Defense Spending Scale |
| VCF0867a | R Opinion: Affirmative Action in Hiring/Promotion [2 of 2] |
| VCF0876a | R Opinion Strength: Law Against Homosexual Discrimination |
| VCF0877a | R Opinion Strength: Favor/Oppose Gays in Military |
| VCF0878 | R Opinion: Should Gays/Lesbians Be Able to Adopt Children |
| VCF0879a | R Opinion: U.S. Immigrants Should Increase/Decrease [2 of 2] |
| VCF0886 | R Opinion: Federal Spending- Poor/Poor People |
| VCF0887 | R Opinion: Federal Spending- Child Care |
| VCF0888 | R Opinion: Federal Spending- Dealing with Crime |
| VCF0889 | R Opinion: Federal Spending- Aids Research/Fight Aids |
| VCF0894 | R Opinion: Federal Spending- Welfare Programs |
| VCF9013 | R Opinion: Society Ensure Equal Opportunity to Succeed |
| VCF9014 | R Opinion: We Have Gone Too Far Pushing Equal Rights |
| VCF9015 | R Opinion: Big Problem that Not Everyone Has Equal Chance |
| VCF9037 | R Opinion: Government Ensure Fair Jobs for Blacks |
| VCF9040 | Blacks Should Not Have Special Favors to Succeed |
| VCF9047 | R Opinion: Federal Spending- Improve/Protect Environment |
| VCF9048 | R Opinion: Federal Spending- Space/Science/Technology |
| VCF9049 | R Opinion: Federal Spending- Social Security |
| VCF9131 | R Opinion: Less Government Better OR Government Do More |
| VCF9132 | R Opinion: Govt Handle Economy OR Free Market Can Handle |
| VCF9133 | R Opinion: Govt Too Involved in Things OR Problems Require |

## 2012 CCES Variables

| Variable | Description |
| --- | --- |
| CC332A | roll-call votes - Ryan Budget Bill |
| CC332B | roll-call votes - Simpson-Bowles Budget Plan |
| CC332C | roll-call votes - Middle Class Tax Cut Act |
| CC332D | roll-call votes - Tax Hike Prevention Act |
| CC332E | roll-call votes - Birth Control Exemption |
| CC332F | roll-call votes - U.S.-Korea Free Trade Agreement |
| CC332G | roll-call votes - Repeal Affordable Care Act |
| CC332H | roll-call votes - Keystone Pipeline |
| CC332I | roll-call votes - Affordable Care Act of 2010 |
| CC332J | roll-call votes - End Don't Ask, Don't Tell |

## Matched roll-call votes, Senate

Roll-call votes are on final passage, where applicable. In the case of issues that were voted on multiple times, we take the vote closest to the 2012 election. Roll call vote data were obtained from `voteview.com`.

## Senate

| Issue | Congress | Vote Number |
| --- | --- | --- |
| Affordable Care Act | 111th | 396 |
| Repeal Don't Ask, Don't Tell | 111th | 281 |
| Tax Hike Prevention Act | 111th | 276 |
| Ryan budget | 112th | 77 |
| Middle Class Tax Cut Act | 112th | 184 |
| US-Korea Free Trade Agreement | 112th | 161 |
| Affordable Care Act Repeal | 112th | 9 |
| Keystone Pipeline | 113th | 280 |

## House of Representatives

| Issue | Congress | Vote Number |
|---|---|---|
| Affordable Care Act | 111th | 165 |
| Repeal Don’t Ask, Don’t Tell | 111th | 317 |
| Tax Hike Prevention Act | 111th | 647 |
| Ryan budget | 112th | 277 |
| Middle Class Tax Cut Act | 112th | 545 |
| US-Korea Free Trade Agreement | 112th | 783 |
| Affordable Care Act Repeal | 112th | 14 |
| Keystone Pipeline | 113th | 519 |

# D Alternative Mixture Model Results

Instead of using the distribution $\Delta_d(x_i)$ for estimating $\pi_d$, we can also use the distribution of $x_i$. It is quickly shown that the average log likelihood of $\pi_d$ based on the mixture of $x_i$ is equal to

$$\frac{1}{N}\sum_{i=1}^{N}\log\left[\pi_d e^{\Delta_d(x_i)} + 1 - \pi_d\right] + \log P_0(x_i), \tag{11}$$

which means we can estimate $\pi_d$ using just the observed values $\Delta_d(x_i)$; no need to approximate $F_0$ or $F_d$, although we still prefer having the ability to visually inspect the mixture model results by using LLR distributions. The results of performing this analysis — using the cross-validated $\hat{\Delta}_d(x_i)$ as in the main analysis — are given in Table D.1. The results are remarkably similar to the results given in Table 2 in the main text. The only noticeable difference is that $\hat{\pi}_1$ for the Public CCES 2012 Roll Calls is significantly higher than the $\hat{\pi}_1$ estimated using the LLR distributions, suggesting that nearly all of the CCES respondent population are ideal point types with a unidimensional ideal point model.

| Population | Data Source | $\hat{\pi}_1$ | $\hat{\pi}_2$ |
|---|---|---|---|
| *Politicians:* | State Legislator Survey | 0.99 | 0.99 |
| | | (0.008) | (0.026) |
| | Roll Call Votes | 0.99 | 0.99 |
| | | (0.036) | (0.036) |
| *Public:* | ANES 2012 | 0.91 | 0.97 |
| | | (0.006) | (0.005) |
| | CCES 2012 Roll Calls | 0.99 | 0.99 |
| | | (0.004) | (0.003) |
| | Broockman (2016) Survey | 0.75 | 0.87 |
| | | (0.026) | (0.020) |

Table D.1: Alternative estimates of $\pi_1$ and $\pi_2$ across data sources, with standard errors shown in parentheses. $\pi_d$ gives the proportion of the population that is better described by a $d$-dimensional ideal point model compared to an unrelated-prefrence model.