Jointly scaling citizens and legislators. The issue of varying structures in ideology

Lukas F. Stoetzer*

[Early draft prepared for EPSA 2016, Brussels]

The representation of citizens and legislators in the same ideological space requires that both employ the same structure to guide their political attitudes. Public opinion research has long argued for a differentiation between the two groups. While previous research assumes that legislators closely align their attitudes along ideological dimensions, citizens are said to do so only to a varying degree. In this manuscript, I describe an item response model with group specific error variances that takes this disparity into account. An application of the model to the cooperative congressional election study (CCES) of 2008-2012 shows that respondents' attitudes about political issues are generally less constrained by underlying dimensions when compared to legislators' preferences revealed in roll call votes on the same issues. The results further highlight that citizens who closely follow the news about public affairs are more similar to legislators in applying ideology to organize their preferences.

*Post-Doc, Department of Political Science, Massachusetts Institute of Technology. E-Mail: lstoetze@mit.edu, Webpage: http://lukas-stoetzer.org/

Many contemporary theories attempting to explain the political process rely on the abstraction of political preferences in terms of a few underlying ideological dimensions. In the spatial model of party competition, political candidates occupy political platforms along the liberal conservative spectrum and citizens use their own ideological orientation to choose among the competing candidates [see e.g. Downs, 1957, Davis et al., 1970, Schofield, 1978, Adams et al., 2005, Laver and Sergenti, 2010]. This parsimonious description of the political process is one of the most influential in contemporary political science and as such, it is hardly surprising that methodological innovations aim to uncover a reliable and valid representation of those basic ideological dimensions. Over the years, the literature has presented methods to estimate the latent positions based on respondents' perceptions of political actors [Aldrich and McKelvey, 1977, Hare et al., 2015], and by scaling actual observed roll call voting behavior in parliaments [Poole and Rosenthal, 1985, Clinton et al., 2004].

More recently, the observed voting approach was employed to jointly scale the ideology of legislators and citizens [Jessee, 2009, Bafumi and Herron, 2010, Shor, 2009]. This work leverages survey questions that ask citizens how they would have voted on a particular bill, thereby bridging citizens' preferences to those of legislators. Researchers employ an item response formulation to recover ideological positions of both groups, creating a joint ideological space. However, discussions of this approach, by Jessee [forthcoming] and Lewis and Tausanovitch [2015], point out that this hinges on the crucial assumption that the item parameters of the model are constant for both groups. Otherwise, the scaled responses do not reflect the same underlying structure. In the words of Lewis and Tausanovitch [2015]: "This constant item (or constant behavior) assumption is the underpinning of joint scaling exercises." If this assumption is unmet, the resulting positions strongly depend on the number of citizens included in the analysis [Jessee, forthcoming]. Both research teams find that the item parameters actually strongly

deviate between the two groups.

One of the conclusions from this research, that the way in which underlying dimensions of politics guide the preferences for policy proposals differs between citizens and legislators, can hardly be described as new. In fact, one of the most prominent models in public opinion and political behavior accompanies this finding with a theoretical foundation. The mass belief system perspective, most famously associated with Phillip Converse's 1964 chapter [see also Feldman and Stanley Feldman, 1988, Jacoby, 1995, Zaller, 1992], contest that the general public is mostly "innocent of ideology" [Kinder, 1983, p.391]. The origin of this model is the idea that political attitudes "are organized into coherent structures by political elites for consumption by the public" [Feldman and Stanley Feldman, 1988, p.417]. To what extent citizens use the same dimensions of politics to align their attitudes strongly depends on the exposure to those structures. If the general public adopts the same ideology considerations as the political elites, ideology should also constrain their opinions about the policies. If not, citizens' attitudes do not mimic the ideological structures. The supposition then is that compared to political elites, the general public does not organize its preferences as strongly along underlying ideological dimensions. Instead, citizens' preferences are most often inconsistent and unstable.

The exercise of creating a common ideological space of citizens and political elites potentially has to take these long-standing arguments about about heterogeneity in ideological constraints seriously. While for legislators lower ideological dimensions are a strong predictor of their policy preferences, for the general public those dimensions might simply be less powerful. Instead, additional unobserved aspects weaken the relationship between ideology and electorates' preferences for policies. To account for this fact, I describe a heterogeneous group-based item response model that takes this variation into account. This model implements the idea that the error variance of factors next to

ideology can be larger for one group compared to another group. It permits for a direct test of the mass belief system perspective. First, citizens' responses to political issues should be more noisy, when compared to those of political elites. Second, there should be considerable heterogeneity within the general public. In particular, individuals who are strongly exposed to public debate should be more similar to political elites than individuals who hardly follow the news.

An application of the model to the cooperative congressional election study (CCES) of 2008-2012 underscores the importance of this line of inquiry. The CCES contains questions that allow me to link respondents' preferences for a particular bill to roll call voting behavior of representatives in the house and the senate. The method is well-suited in uncovering systematic differences in the answering patterns. The results confirm the supposition that respondents' answers are generally more noisy as opposed to legislative roll call votes. The method can further be deployed to pinpoint differences within the general public. As such, I find that respondents who closely follow current affairs, more strongly align their preferences in line with underlying dimensions, compared to those who only hardly follow the news. This means that citizens who are strongly exposed to the ideological structure are more similar to legislators in applying ideology to organize their preferences. The manuscript closes with a discussion on further developments of this method.

1. An item response model with group specific error variance

In this section, I describe an item response model that allows for group specific error variance. Standard IRT models assume that the response to a particular item are systematically influenced by underlying traits and that each response experiences a stochastic

deviation from the trait-expected pattern. In the application of IRT to roll call votes, preferences for a particular policy can be understood in terms of a legislators' latent orientation. A liberal senator is more likely to be in favor of a minimum wage, but on each proposal she might deviate from her ideological expected vote. While it is commonly assumed that the variance of this stochastic part is constant for all legislators (and as such is used in the application of joint scaling), this assumption can be generally be relaxed. For parts of the respondents the positions on the underlying dimensions might be less predictive, implying that the stochastic variation is larger. This idea is especially intriguing when analyzing differences between two groups, e.g. legislators and citizens. Respondents who express their preferences for a particular bill in a survey might exhibit a larger stochastic variance in relation to legislators' voting behaviour in parliament.

The model can be formalized as follows: Suppose each individual i ($i \in \{1, \ldots, N\}$) can be assigned to one of the distinct groups $g(i) \in \{1, \ldots, G\}$ and all answer the same j ($j \in \{1, \ldots, J\}$) items. There is a different number of individuals in each group N_g. I assume that the response to item j follows the same logic as outlined in spatial voting in parliament, which is analogous to a two-parameter IRT Model [Clinton et al., 2004]. However, in my application the error variance is allowed to vary between grops. The answer to item $y_{ij} \in \{0, 1\}$ is positive if $y_{ij} = \beta_j \theta_i - \alpha_j + \epsilon_{ij}$, where the errors are distributed i.i.d. normal $\epsilon_{ij} \sim N(0, \sigma_{g(i)})$ with group specific variance $\sigma_{g(i)}$. In this set-up, β_j correspondents to the *discrimination*, α_j to the *difficulty* parameter and θ_i to the individual *trait*[1]. Based on this, the probability of correct response is given by:

$$Pr[y_{ij} = 1] = \Phi \left[\frac{\beta_j \theta_i - \alpha_j}{\sigma_{g(i)}} \right], \tag{1}$$

[1]Throughout the article, I work with the assumption that the traits are one-dimensional. Data limitations in the application to the CCES do not allow for a more dimensional representation of the traits.

where Φ is the C.D.F of the standard normal distribution. $\sigma_{g(i)}$ scales the systematic impact of the latent trait. Higher values of $\sigma_{g(i)}$, decrease the impact θ_i has on the predicted probability to answer a specific item positively. If $\sigma_{g(i)}$ is one for all groups the the formulation corespondents to the standard two parameter item response model.

The description of the item response model with group specific variance originates from a general heteroskedastic IRT model [Lauderdale, 2010], in which each respondent can stochastically deviate from the expected answering patterns to a different degree. In Lauderdale's model, each respondent has an idiosyncratic error variance $\epsilon_{ij} \sim N(0, \sigma_i)$. It is possible to identify this variance because each respondent answers multiple items. But an accurate estimate of those individual error variances requires plenty of items, which are in most applications not available. The application here is instead tailored towards comparing variation between groups. Although this is aimed at investigating differences between legislators and citizens, the model can generally be employed to analyze the predictive power of the latent trait for any clustering of respondent groups. This is of particular interest to this empirical application as it allows me investigate if politically interested respondents are more like representatives.

It is well understood that the model as presented above is not identified [Rivers, 2003]. The one-dimensional version of the model can locally be identified by fixing the location and scale of the latent dimension. I achieve this using appropriate prior specifications for the parameters and transformation of some parameters. Priors for the θ_i are standard normal distributed and all priors on the item parameters (α_j, β_j) are specified to be normal around zero with a standard deviation of one. Priors on the error variances are inverse gamma distributions $\sigma_{g(i)} \sim IG(100, 100)$. I choose priors that are slightly informative following the belief that there should be no difference between groups. Moreover, to identify the scale, I follow Lauderdale [2010] and constrain the mean of the inverse group specific error variance to be equal to one $\frac{1}{G}\sum \frac{1}{\sigma_g} = 1$. The

traits (θ_i) are also normalized with mean zero and standard deviation of one. The two constraints are implemented by rescaling the values at the each iteration of the Gibbs sampler routine.

MCMC routines to sample from the posteriori of the two parameter IRT model are discussed in Clinton et al. [2004], Martin and Quinn [2001] and are implemented in R the package *pscl* [Jackman, 2008]. For the item response model with group-specific variance, I implement my own program relying on RcppArmadillo [Francois et al., 2012]. My implementation closely follows the gibbs sampler discussed in Lauderdale [2010], with the exception that the error variances are not sampled for each respondent, but for each group. Only the last step differs compared to Gibbs sampler for the heteroskedastic IRT model. Instead of calculating the sum of squared errors (on the latent sale) for each individual they are calculated for the group SSE_g. The group specific error variances are then drawn from the inverse gamma conditional distribution $\pi(\sigma_g \mid y^*_{ij}, \alpha_j, \beta_j, \theta_i) \sim IG\left(\frac{100+N_gJ}{2}, \frac{100+SSE_g}{2}\right)$. This step has a straightforward interpretation. With an increasing sum of squared errors for a group (meaning that the model fits the responses for a particular group less well) the expectation of the error variance for this group also increases. The error variances, hence, picks up how well the IRT structure fits for each group.

2. Application: CCES 2008-12 roll call vote questions

This section describes an application of the model to the cooperative congressional election study (CCES) of 2008, 2010 and 2012 [Ansolabehere, 2013, 2012, Ansolabehere Stephen; Schaffner, 2015]. All three studies include questions that can be linked to roll call votes in the house and/or senate. For example, the 2008 CCES asks respondents about their preferences for nine political issues. The issues are introduced as: "Congress considered many important bills over the past two years. For each of the following tell us

whether you support or oppose the legislation in principle.". For each of the items, e.g. "withdraw Troops from Iraq within 180 days", respondents can indicate if they support this bill, oppose it, or are "not sure". The introductory question text in the 2010 and 2012 CCES is the same, but the items differ[2]. The items cover a wide range of topics such as economic issues (minimum wage, taxation, bank bailout, budget cuts) and social issues (gay marriage, end don't ask don't tell act, birth control). For a complete list of items included in the studies please refer to appendix A. The items match roll call votes in congress [Pettigrew, 2013a,b,c].

In the following analysis, the matching roll call votes are employed to capture legislators' preferences for those items. Each CCES study generally correspondent to one congress (CCES 08 to 110th congress, CCES 10 to 111th congress and CCES 12 to the 112th congress)[3]. House and senate roll call votes on the issues are pooled together, giving around 545 legislators who vote on the issues for each of the corresponding studies[4]. A one is coded as a favorable vote and a zero as a unfavorable vote. The expressed opinions in the CCES studies are taken as indication of citizens' preferences, coding ones as a positive attitude towards the issue and zero as a negative. "Not sure" answers are coded as missing values. In total, this creates two response matrices (one for citizens, and one for legislators) for each of the studies. The responses can be used to scale ideological convictions, and to analyze how the revealed preferences relate to one underlying dimension.

[2]Some items overlap, which in principal would make it possible to pool the studies.

[3]The CCES questionnaire also includes roll call vote questions from the previous senate. E.g. the CCES 2010 asks respondents about their attitude towards the troubled asset relief program, a roll call vote of the 110th senate. In this case, the votes of the senators and house members are take from the previous congress. The decision of newly elected representatives on those issues are coded as missing values.

[4]For some bills roll call votes are only available for either the house or the senate. For the five items in the CCES 2012 were this is the case, all non-existing house (or senate) votes are coded as missing

2.1. Respondents' and legislators' discrimination parameters

The first step of the analysis is to see how the structure in ideology differs between legislators and citizens. If the responses can be traced back to the same underlying construct, the item parameters of the item response model should be similar. Discrepancies in parameter estimates can be illuminating for the disparity between citizens in legislators. Estimating the ideal model of roll call votes [Clinton et al., 2004] for both groups and all studies separately, allows contrasting juxtaposition of the item parameters[5]. The model is estimated for the entire legislator response matrix. For the citizens, a random subset of all respondents in the CCES studies is drawn that is of the same size as the response matrix of the legislators. This approach can be though of as an experiment, asking how the same number of citizens would have voted on those proposals and if their decisions can be reduced to underlying dimensions in a similar way[6].

Figure 1 contrasts the discrimination parameters obtained for the legislators with parameters estimated based on respondents' answers for the different CCES studies[7]. The y-axis represents median discrimination draws for legislators, and the x-axis the median draw for the respondents. The first aspect to note is that most parameters are positive or negative for both groups, implying that they discriminate in the same direction. E.g. citizens that favor withdrawing troops from Iraq are more liberal, as are representatives who are in favor of this legislation. One exception is the Tax Hike prevention act in the CCES 2010. Apart from this issue, all items show the same signs. Some are rather

[5] All models in this section are estimated using ideal from the *pscl* R-package [Jackman, 2008], with 50'000 burn-in draws and 200'000 draws, saving every 100th observation. Starting values are set for legislators according to their party affiliation (Democrats -1 and Republican 1). For the respondents , I use their reported party identification as starting values (Strong Democrats at -1, strong republicans at 1 and all other at zero). The starting values make convergence to the same posteriori mode likely.

[6] Another reason for working with sub-samples is to ease computation. With up-to 50'000 responses in the CCES the Bayesian estimation would take up considerable amount of time and memory, while not necessarily changing the conclusions drawn about the item parameters. Multiple estimations of the model on different sub-sets only exhibit minimal variation.

[7] The discussion here is focused on the more insightful discrimination parameters. For the corresponding plot of the difficulty parameters please refer to the appendix C figure 5.

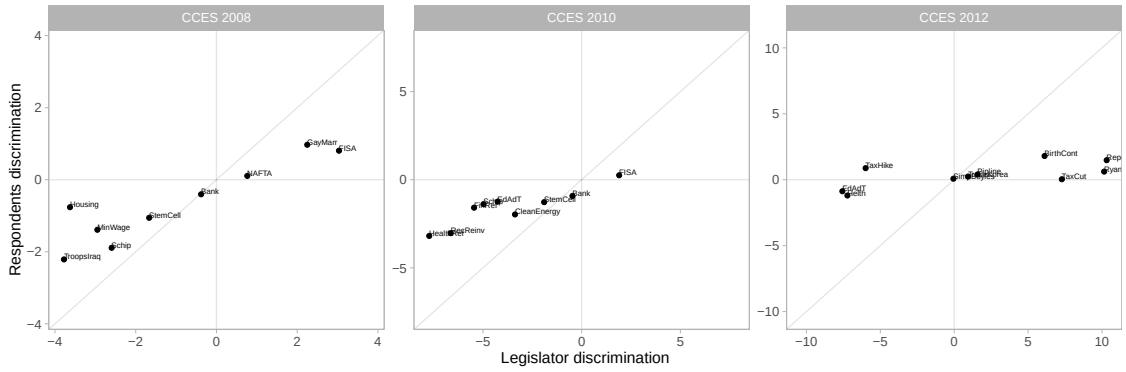

Figure 1: Comparison of discrimination parameters between legislator scaling and respondent scaling in the CCES 2008-12. The legislator scaling is based on house as well as senator legislators. The respondent scaling is based on a random sub-sample of the same size as the legislators.

close to zero for the citizens. For example, the credible interval for NAFTA ("Extend the North American free Trade agreement to include Peru and Columbia") in the CCES 2008 includes zero, thereby it does not strongly differentiate between liberal and conservative respondents, but it does so (although mildly) for legislators. Another intriguing fact is that the ordering of the strength in discrimination is highly correlated. Especially, in the CCES 2008 and the CCES 2010 the item ranks (according to their discrimination) for legislators and citizens are similar. TroopsIraq has the strongest negative discrimination for citizens and legislators in CCES 2008 and so does the Health Reform issue in the CCES 2010. For the CCES 2012 the discrimination ranking is arguably more messy. Nonetheless, the relationships reveal that underlying traits structure preferences in the same direction, and the issue preferences appear to origin from a similar dimension.

Of particular interest is the 45 degree line, as it represents a one-to-one relationship. Item parameters that fall on this line would indicate the exact same discrimination. Observations that fall above (for negative values) or below (for positive values) (but below/above the zero line) present weaker discrimination for respondents compared to legislators. It is telling that almost all observations fall in this area. Items do not as sharply differentiate liberal from conservative respondents, as they do for legislators.

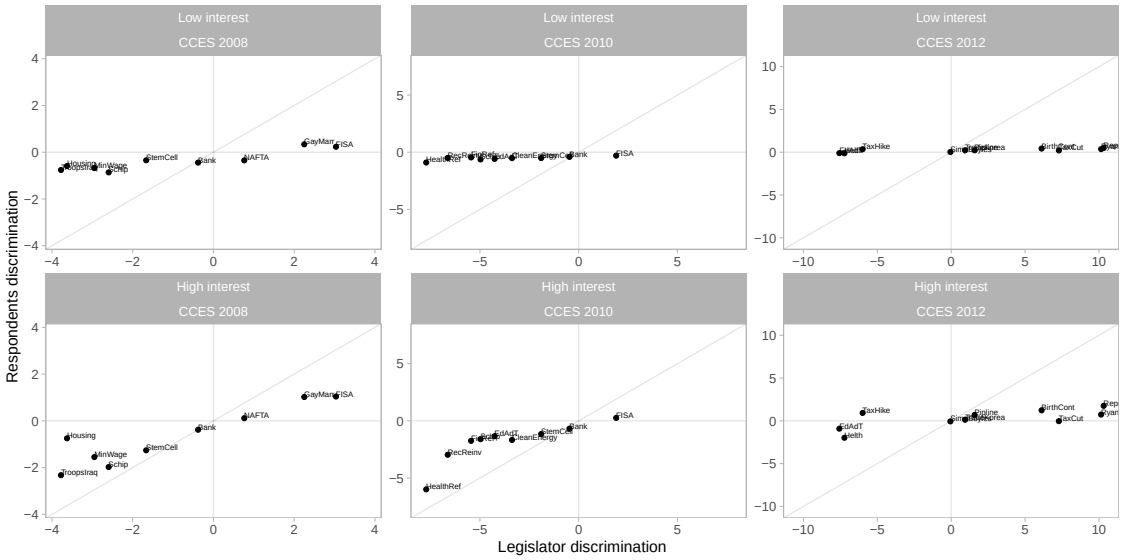

Figure 2: Comparison of discrimination parameters between legislator scaling and respondent scaling in the CCES 2008-12 for different levels of political interest. The respondent scaling is based on a random sample of political interested and none-interested respondents.

This implies that responses on the survey questions can be described as more noisy. For the CCES 2008 and the CCES 2010 the relationship appears to be systematic. All points scatter around a lower degree line. A pattern like this can be explained by the item response model with group specific error variance (in equation 1). It emerges if for respondents all discrimination parameters are divided by a error variance that is larger in relation to the one of legislators. However, as evidence for the mass-belief perspective, and a lower degree of ideological constraint within the electorate, the results are not sufficient. After all, the increased error variance can originate from the measurement instrument itself. The disparity between voting on a lengthy legislation and a one-sentence survey question is obvious. Survey responses, as uncostly expressions of thoughts on the top-of the respondent's head, are likely to examine stronger error variance compared to observable and costly roll call votes in parliament.

An additional investigation reveals difference within the general public. Measurement error that originates from the nature of the survey instrument should be constant for

all respondents. If some specific parts of the respondent pool are closer to the pattern of the legislators, whereas other parts are further away, this would hint at the existence of varying ideological constraints. To test this idea, I re-estimated each of the models twice. Once using a random sample of respondents who indicate that they "follow whats going on in government and public affairs" most or some of the time, and once for respondents who follow the news only now and then or hardly at all[8]. Political interest and information should be a reliable predictor of the variation in ideological structure. The first group should be closer to the legislators, compared to the second politically uninterested group.

Figure 2 displays the results. The top three panels show the comparison for the sample of respondents that has a low interest in public affairs. The lower three panels depict the relationship for the group with a strong interest in public affairs. It is evident that for the low interest group the discrimination parameters are systematically smaller compared to the legislators, while for the high interest group the pattern closer mimics the one found for legislators. For the CCES 2008 the patterns most strongly align with the expectation. In four out of nine items, the credible interval of the discrimination parameter for legislators overlap with those of the high interest group. For the low interest group three parameters cannot be distinguished from zero. Similar results hold for the CCES 2010 and the CCES 2012. This yields some support for the mass-belief system perspective. Politically uninterested voters do not organize their preferences along ideology as strongly as political elites. For well-informed and political interested parts of electorate, preferences are guided by ideology in similarly strong way as those of political elites. This would also imply that for this part of the electorate the constant parameter assumption is mostly meet. The next section discusses in how far similar patterns can be picked-up using the item response model with group specific variances.

[8]The survey question is included in all CCES studies and thereby makes a direct comparison possible. Most respondents are found in the highest interest category. Only a small subset reports that they hardly follow the news at all. For descriptive statistics please refer to appendix B.

	CCES 2008		CCES 2010		CCES 2012	
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Legislators (σ_1)	0.73	0.7	0.9	0.83	0.71	0.74
	[0.71,0.76]	[0.66,0.73]	[0.85,0.95]	[0.78,0.88]	[0.68,0.75]	[0.7,0.79]
Citizens (σ_2)	1.57		1.13		1.67	
	[1.47,1.7]		[1.05,1.21]		[1.52,1.86]	
Int. citizens (σ_2)		1.24		0.93		1.1
		[1.16,1.33]		[0.88,0.99]		[1.03,1.18]
No int. citizens (σ_3)		1.32		1.39		1.35
		[1.23,1.43]		[1.28,1.51]		[1.25,1.47]
Items	9	9	9	9	10	10
Legislators	547	547	541	541	544	541
Citizens	547	1094	541	1082	544	1088

Table 1: Parameter estimates for item response model with group specific error variance.

2.2. Results from item response model with group specific error variance

The item response model with group specific variance is estimated for the three studies. The first type of specifications include two groups: citizens and legislators. For the citizens, a random subset of respondents (of the same size as the number of legislators) is included in the analysis. This means that the analysis gives equal weight to respondent and legislator observations[9]. For the second type of specifications, three groups are created: Legislators, a group of political interested citizens, and a group of politically uninterested citizens. All groups are of equal size. For the citizen groups, respondents are differentiated based on their answer to the follow the news question. The same question that has been used in the analysis above. The political interested citizens group contains of a random sample of respondents who follow the news most or some of the time. For the other group, a random sample is created from respondents who only sometimes or hardly at all follow public affairs[10]. The models are estimated using the sampling scheme discussed in section above. It is run for 250'000 iteration, discarding

[9] In how far this influences the results is critically investigated in the discussion.

[10] For the CCES 2010 the number of respondents for this group is actually not large enough to draw a random sample of the same size than the legislators. In this case, I use re-sampling of observations.

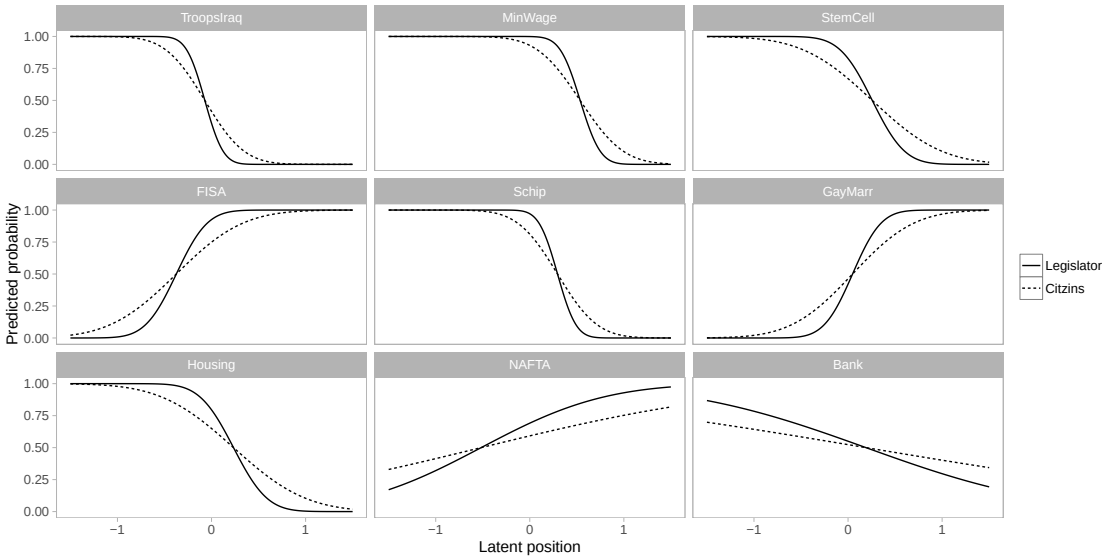

Figure 3: Item response curves for legislators and citzins for different items from the CCES 2008. Estimated from item response model with group specific error variance (Model 1).

the first 50'000 as burnin and saving every 100 draw for analysis of the posterior[11].

Table 2.2 reports the results for the estimated group specific error variances[12]. In all three studies the estimated error variance is larger for the citizens compared to the legislators. The strongest difference is observed in the CCES 2012 and CCES 2008. In the CCES 2008, the median of the posteriori error variance is 0.73 for legislators and 1.57 for citizens. This has a strong effect on the way how variation in ideological orientations influences preferences for the issues. Figure 3 plots the item response curves. The straight lines are the item response curves for the legislators, the dotted lines for the citizens. Because the estimated error variance is larger, the slopes of the item response curves are lower for citizens. This induces more randomness in their responses. Similar patterns emerge for the CCES 2012 with median of the posteriori error variance of 0.71 for legislators and 1.67 for citizens. Figure 8 (attached in the appendix D) discloses

[11]For a comparison of ideal point estimates for the legislators to the once obtained above see appendix D figure 10. The median estimates are highly correlated. Only for the CCES 2010 the relationship is quite noisy.

[12]The complete parameter estimates can be found in the appendix D. Figure 6 shows the discrimination as well as difficulty parameters. Figure 9 the distribution of latent traits.

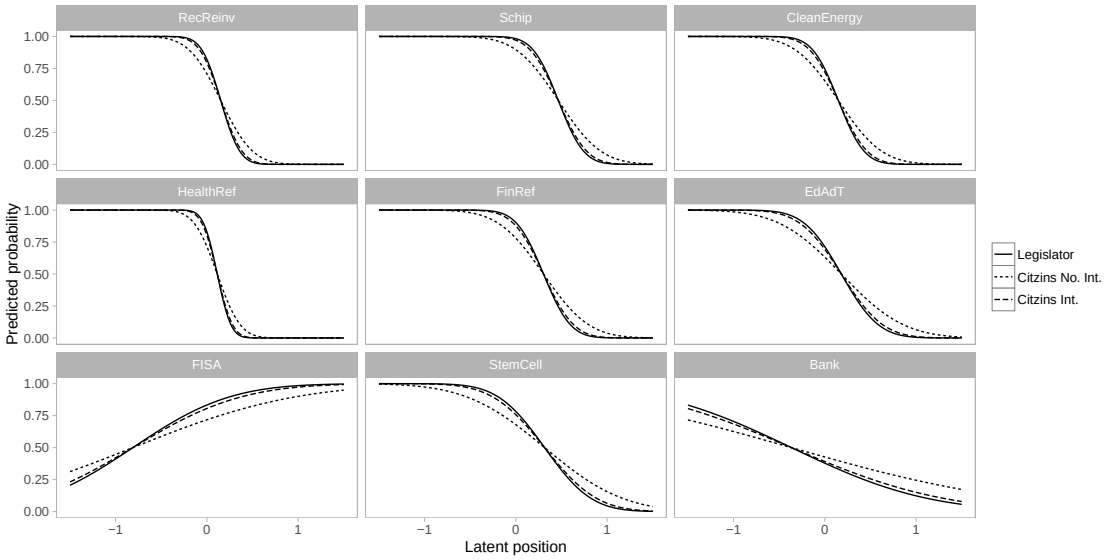

Figure 4: Item response curves for legislators and citzins groups for different items from the CCES 2010. Estimated from item response model with group specific error variance (Model 3).

similar differences in item response curves between Representatives and respondents. However, in this case, some of the item exhibit little discriminative power for neither of the groups. In the CCES 2010 the difference in error variance is not as distinct (0.9 for legislators and 1.13 for citizens), which results in item response curves that are almost coinciding (see figure 7 appendix D). All in all, the models permit to pinpoint the difference between responses on the survey items and legislators roll call vote behavior in terms of increasing error variance.

Further dividing citizens in two political interest groups highlights that the answers of the political interested part of the general public is more similar to legislators. While in the CCES 2008 the posteriori distributions of the error variance do not strongly differ between the citizens groups (1.32 and 1.24), for the other studies the informed group's estimated error variance is closer to one of the legislators. Especially, in the CCES 2010 (Model 4), the error variances for the informed part is very similar to the legislators (0.83 and 0.99). For the CCES 2012 the difference between the citizen groups is present

(1.35 and 1.1), but legislators are still by far the group with the lowest error variance. Figure 4 illustrates the result for the CCES 2010. The item response curves are similar for legislators and the respondent sample that exposes itself to the public affairs. For the group that only sometimes or hardly follows the news, the variation in underlying orientations less strongly constraint preferences for the items. This illustrates how the method can be deployed to uncover which parts of the public most closely mimic the item response structure of the political elites. It further supports the idea of the mass belief system perspective that the exposure to the ideological structure, transmitted via the discussion of public affair, makes respondents preferences more ideological organized.

3. Discussion

This manuscript discusses an item response model with group specific variance to jointly scale citizens and legislators. In an application to the CCES the model reveals interesting difference between the two groups. In line with argumentations of the mass belief system literature, respondents answering patterns are more noisy and appear less constraint by underlying dimensions of politics, compared to the preferences of legislators. Additionally, within the electorate those who expose themselves regular to news about public affairs more strongly follow the attitude structure found in roll call votes in parliament. This discussion is dedicated to further developments of the model.

The first aspect that remains undisclosed in the main text, is in how far the method helps to overcome some of the issues associated with joint scaling. If the entire difference between citizens and legislators can be attributed to the increased error variance, it might help to solve some of the challenges. It potentially would recover the unmet constant behavior assumption, by including an additional factor that describes difference between groups. However, a first test of this idea fails. With an increasing number of respondents included in the analysis, the method estimates a decreasing difference in error variances

for the two groups. In this, the method is not able to pass the test advocated by Jessee [forthcoming]. For a description and discussion of the results please see appendix E.

This has important implications for further developments of the method. If the main aim of the method is to test which part of the electorate shares similar structures in ideology compared the political elites, it would make sense to estimate the item parameters only based on responses of the latter group and investigate which parts of the other group most closely follow the same structure. This procedure would conform to the conclusions drawn in Jessee's forthcoming contribution. Moreover, it would be useful to not arbitrarily split responses into different groups, but rather use a set of covariates to model the variation in error variances. An extension in this direction, would allow to analyze further aspects as well. E.g. are partisan responses similarly structured than representatives' preferences? Additionally, as the focus is not any particular point in time, it might be sensible to pool information from the different studies. On that note, the first results presented in this manuscript are rather understood as a motivation to further investigate the subject and develop methods that can crystallize how ideology structures political preferences differently for political elites and citizens.

References

J Adams, S Merrill, and B Grofman. *A unified theory of party competition: A cross-national analysis integrating spatial and behavioral factors*. Cambridge Univ Pr, 2005.

John H Aldrich and Richard D McKelvey. A Method of Scaling with Applications to the 1968 and 1972 Presidential Elections. *The American Political Science Review*, 71 (1):111–130, 1977.

Stephen Ansolabehere. CCES Common Content, 2010, 2012. URL `http://hdl.handle.net/1902.1/17705`.

Stephen Ansolabehere. CCES, Common Content, 2008, 2013. URL `http://hdl.handle.net/1902.1/14003`.

Brian Ansolabehere Stephen; Schaffner. CCES Common Content, 2012, 2015. URL `http://hdl.handle.net/1902.1/21447`.

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

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

P E Converse. *The Nature of Belief Systems in the Mass Public*. New York, Free Press, 1964.

Otto A Davis, Melvin J Hinich, and Peter C Ordeshook. An expository development of a mathematical model of the electoral process. *American Political Science Review*, 64 (2):426–448, 1970.

Anthony Downs. *An economic theory of democracy*. Harper & Brothers, 1957.

Stanley Feldman and Stanley Feldman. Structure and Consistency in Public Opinion: the Role of Core Beliefs and Values. *American Journal of Political Science*, 32(2): 416–440, 1988.

Romain Francois, Dirk Eddelbuettel, and Doug Bates. RcppArmadillo: Rcpp integration for Armadillo templated linear algebra library. *R package version 0.3*, 6, 2012.

C Hare, D A Armstrong, R Bakker, R Carroll, and K T Poole. Using Bayesian Aldrich-McKelvey Scaling to Study Citizens' Ideological Preferences and Perceptions. *American Journal of Political Science*, 59(3):759–774, 2015.

Simon Jackman. pscl: classes and methods for R developed in the political science computational laboratory, 2008.

William G Jacoby. The structure of ideological thinking in the American electorate. 1995.

Stephen A Jessee. Spatial voting in the 2004 presidential election. *American Political Science Review*, 103(1):59–81, feb 2009.

Stephen A Jessee. (How) can we estimate the ideology of citzins and political elites on the same scale? *American Journal of Political Science*, forthcoming.

D R Kinder. Diversity and complexity in american public opinion. In Ada W Finifter, editor, *Political science: The state of the discipline*, pages 389–425. The American Political Science Association, 1983.

Benjamin E. Lauderdale. Unpredictable voters in ideal point estimation. *Political Analysis*, 18(2):151–171, 2010.

M Laver and E Sergenti. Party Competition: an Agent-Based Model, 2010.

Jeffrey B Lewis and Chris Tausanovitch. When Does Joint Scaling Allow For Direct Comparisons of Preferences? 2015.

Andrew D Martin and Kevin M Quinn. Dynamic Ideal Point Estimation via Markov Chain Monte Carlo for. pages 1–37, 2001.

Stephen Pettigrew. Roll Call Votes - 110th Congress- 2008 CCES, 2013a. URL `http://hdl.handle.net/1902.1/20537`.

Stephen Pettigrew. Roll Call Votes - 111th Congress- 2010 CCES, 2013b. URL `http://hdl.handle.net/1902.1/20539`.

Stephen Pettigrew. Roll Call Votes - 112th Congress- 2012 CCES, 2013c. URL `http://hdl.handle.net/1902.1/20540`.

Keith T. Poole and Howard Rosenthal. A Spatial Model for Legislative Roll Call Analysis. *American Journal of Political Science*, 29(2):357–384, may 1985.

Douglas Rivers. Identification of multidimensional spatial voting models. Technical Report: Political Methodology Working Papers., 2003.

Norman Schofield. Instability of simple dynamic games. *Review of Economic Studies*, 45(3):575–594, 1978.

B Shor. All together now: Putting congress, state legislatures, and individuals in a common ideological space. In *Ideal Point Modeling Conference, MIT May 2015*, 2009.

John R Zaller. *The nature and origins of mass opinion*. University of Cambridge Press, 1992.

A. Description of Roll Call Vote Data from CCES 2008-2012

ID	CCES	Short Description	CCES Wording
rv1	CC316a	Withdraw Troops Iraq	Withdraw Troops from Iraq within 180 days
rv2	CC316b	Increase Minimum Wage	Increase Minimum Wage from $5.15 to $7.25
rv3	CC316c	Stem Cell Research	Allow federal funding of embryonic stem cell research
rv4	CC316d	Foreign Intelligence Surveillance Act	Allow U. S. spy agencies to eavesdrop on overseas terrorist suspects without first getting a court order
rv5	CC316e	Children's Health Insurance	Fund a $20 billion program to provide health insurance for children in families earning less that $43,000
rv6	CC316f	Anti Gay Marriage Amendment	Constitutional Amendment banning Gay Marriage
rv7	CC316g	Federal Assistance for Housing Crisis	Federal assistance for homeowners facing foreclosure and large lending institutions at risk of failing
rv8	CC316h	Extend NAFTA	Extend the North American Free trade Agreement (NAFTA) to include Peru and Columbia
rv9	CC316i	Bank Bailout	Governments $700 Billion Bank Bailout Plan

Table 2: CCES 2008 Roll Call Vote Questions

ID	CCES	Short Description	CCES Wording
rv1	CC332A	Recovery and Reinvestment	Authorizes $787 billion in federal spending to stimulate economic growth in the U.S.
rv2	CC332B	Children's Health Insurance	Program insures children in low income households. Act would renew the program through 2014 and include 4 million additional children.
rv3	CC332C	Clean Energy	Imposes a cap on carbon emissions and allows companies to trade allowances for carbon emissions. Funds research on renewable energy.
rv4	CC332D	Health Reform	Requires all Americans to obtain health insurance. Allows people to keep current provider. Sets up health insurance option for those without coverage. Increase taxes on those making more than $280,000 a year
rv5	CC332E	U.S. Supreme Court	Appoint Elena Kagan to the U.S. Supreme Court
rv6	CC332F	Financial Reform Bill	Protects consumers against abusive lending. Regulates high risk investments known as derivatives. Allows government to shut down failing financial institutions.
rv7	CC332G	End Dont Ask, Dont Tell	Would allow gays to serve openly in the armed services
rv8	CC332H	Foreign Intelligence Surveillance Act	Allow U.S. spy agencies to eavesdrop on overseas terrorist suspects without first getting a court order
rv9	CC332I	Embryonic Stem Cell Research	Allow federal funding of embryonic stem cell research
rv10	CC332J	Troubled Asset Relief Program	$700 billion loans to banks to stabilize finance

Table 3: CCES 2010 Roll Call Vote Questions

ID	CCES	Short Description	CCES Wording
rv1	CC332A	Ryan Budget Bill	The Budget plan would cut Medicare and Medicaid by 42%. Would reduce debt by 16% by 2020.
rv2	CC332B	Simpson-Bowles Budget Plan	Plan would make 15% cuts across the board in Social Security, Medicare, Medicaid, and Defense, as well as other programs. Eliminate many tax breaks for individuals and corporations. Would reduce debt by 21% by 2020.
rv3	CC332C	Middle Class Tax Cut Act	Would extend Bush era tax cuts for incomes below $200,000. Would increase the budget deficit by an estimated $250 billion
rv4	CC332D	Tax Hike Prevention Act	Would extend Bush-era tax cuts for all indviduals, regardlessrv19 of income. Would increase the budget deficit by an estimated $405 billion.
rv5	CC332E	Birth Control Exemption	A Bill to let employers and insurers refuse to cover birth control and other health services that violate their religious beliefs.
rv6	CC332F	U.S.-Korea Free Trade Agreement	Would remove tariffs on imports and exports between South Korea and the U.S.
rv7	CC332G	Repeal Affordable Care Act.	Would repeal the Affordable Care Act.
rv8	CC332H	Keystone Pipeline	A bill to approve the Keystone XL pipeline from Montana to Texas and provide for environmental protection and government oversight.
rv9	CC332I	Affordable Care Act of 2010	Requires all Americans to obtain health insurance. Allows people to keep current provider. Sets up health insurance option for those without coverage. Increases taxes on those making more than $280,000 a year.
rv10	CC332J	End Dont Ask, Dont Tell	End Dont Ask, Dont Tell. Would allow gays to serve openly in the armed services..

Table 4: CCES 2012 Roll Call Vote Questions

B. Descriptive statistics CCES

Table 5: Descriptive Statistics CCES 2008

Statistic	N	Mean	St. Dev.	Min	Max
rv1	28,582	0.534	0.499	0	1
rv2	30,542	0.773	0.419	0	1
rv3	27,440	0.637	0.481	0	1
rv4	28,173	0.684	0.465	0	1
rv5	27,648	0.686	0.464	0	1
rv6	29,277	0.476	0.499	0	1
rv7	25,561	0.498	0.500	0	1
rv8	21,245	0.475	0.499	0	1
rv9	24,155	0.268	0.443	0	1
polint_news	32,367	3.496	0.787	1	4

Table 6: Descriptive Statistics CCES 2010

Statistic	N	Mean	St. Dev.	Min	Max
rv1	54,683	0.480	0.500	0	1
rv2	54,924	0.689	0.463	0	1
rv3	54,341	0.536	0.499	0	1
rv4	54,979	0.489	0.500	0	1
rv5	53,409	0.483	0.500	0	1
rv6	54,429	0.669	0.471	0	1
rv7	54,456	0.580	0.494	0	1
rv8	7,544	0.758	0.428	0	1
rv9	54,666	0.631	0.482	0	1
rv10	7,617	0.268	0.443	0	1
polint_news	54,440	3.531	0.797	1	4

Table 7: Descriptive Statistics CCES 2012

Statistic	N	Mean	St. Dev.	Min	Max
rv1	53,239	0.187	0.390	0	1
rv2	53,656	0.487	0.500	0	1
rv3	51,820	0.614	0.487	0	1
rv4	51,754	0.267	0.442	0	1
rv5	53,340	0.392	0.488	0	1
rv6	52,264	0.509	0.500	0	1
rv7	52,171	0.446	0.497	0	1
rv8	48,757	0.734	0.442	0	1
rv9	53,765	0.568	0.495	0	1
rv10	53,688	0.677	0.468	0	1
polint_news	53,307	3.319	0.906	1	4

C. Results from scaling respondents and legislators

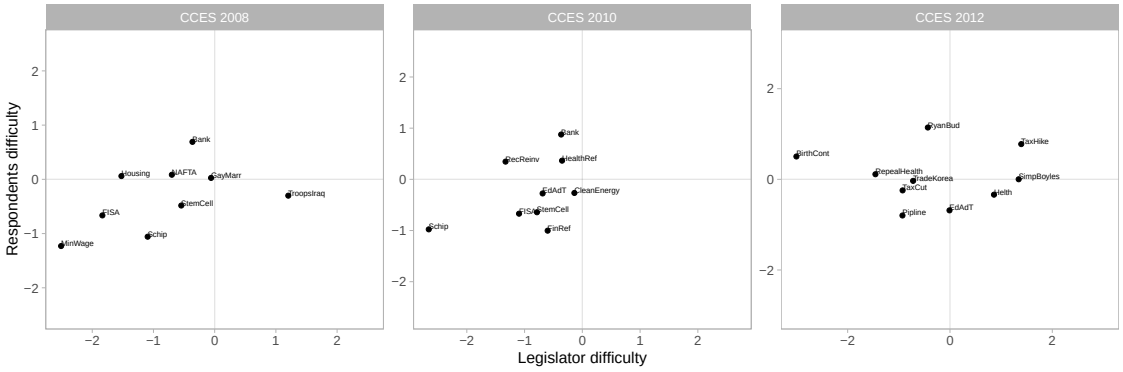

Figure 5: Comparison of difficulty parameters between legislator scaling and respondent scaling in the CCES 2008-12. The legislator scaling is based on house as well as senator legislators. The respondent scaling is based on one random sample of same size as the legislators.

D. Results from item response model with group specific error variance

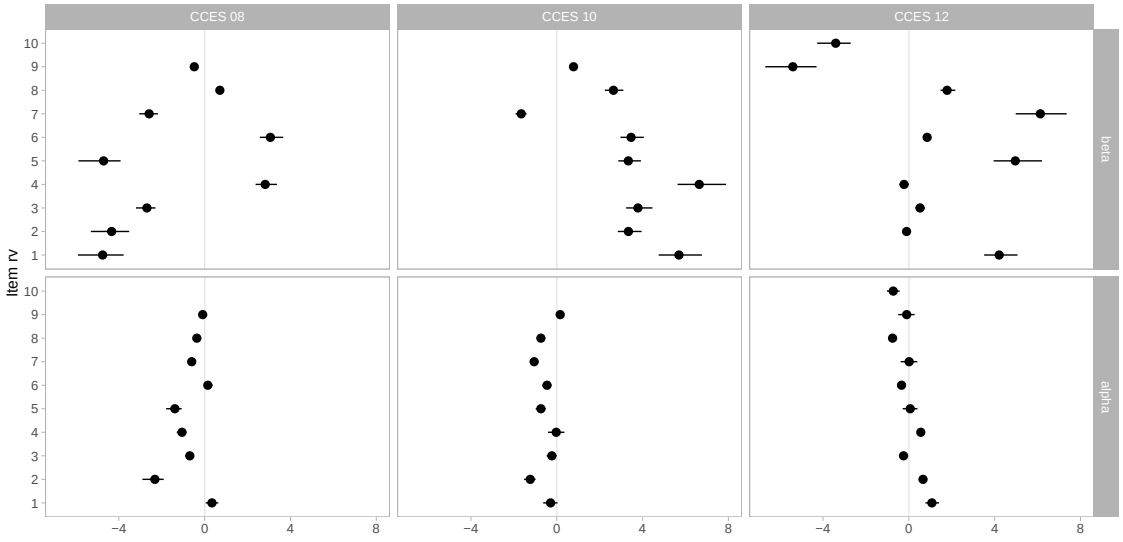

Figure 6: Discrimintaion and difficulty paramter estimates from item response model with group specific error variance

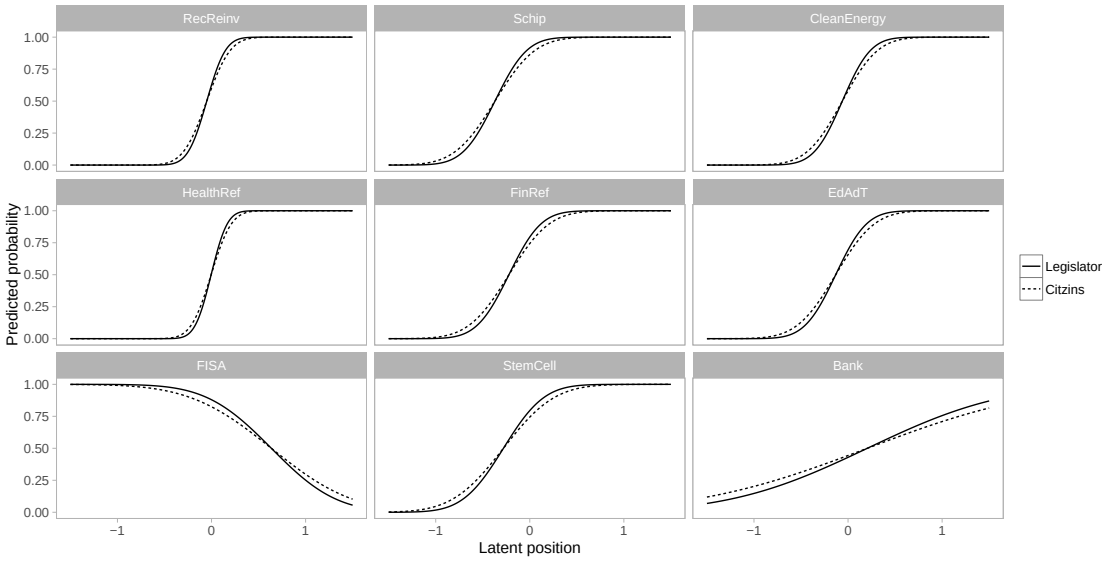

Figure 7: Item response curves for legislators and citzins for different items from the CCES 2010. Estimated from item response model with group specific error variance in model 3.

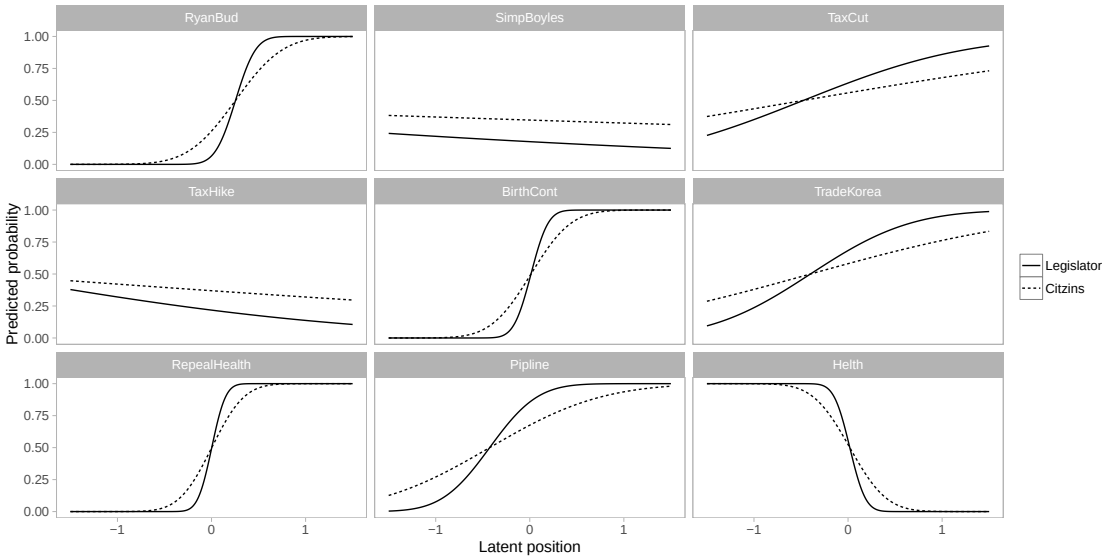

Figure 8: Item response curves for legislators and citzins for different items from the CCES 2012. Estimated from item response model with group specific error variance in model 5.

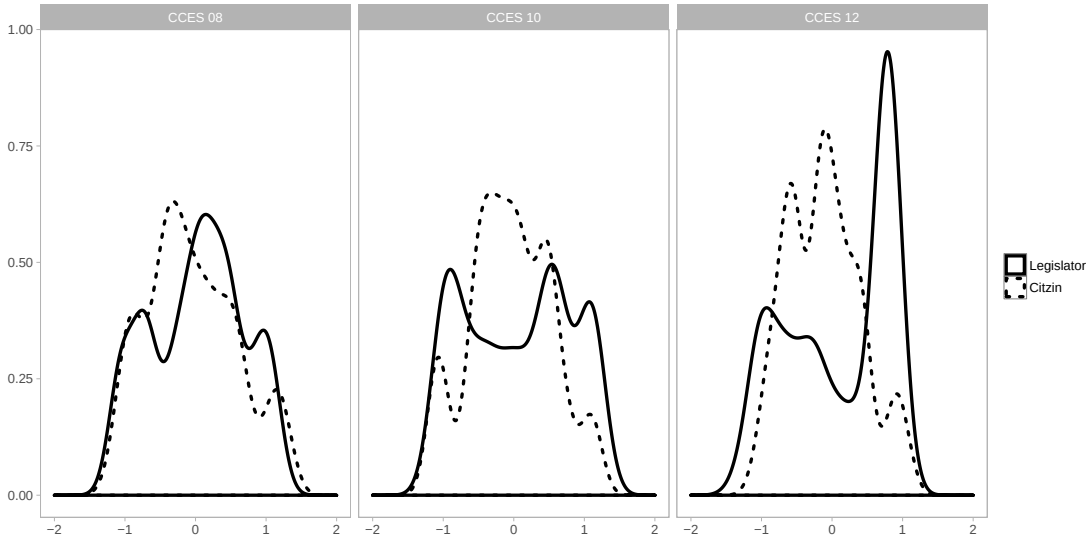

Figure 9: Distribution of latent traits calculated from item response model with group specific variance

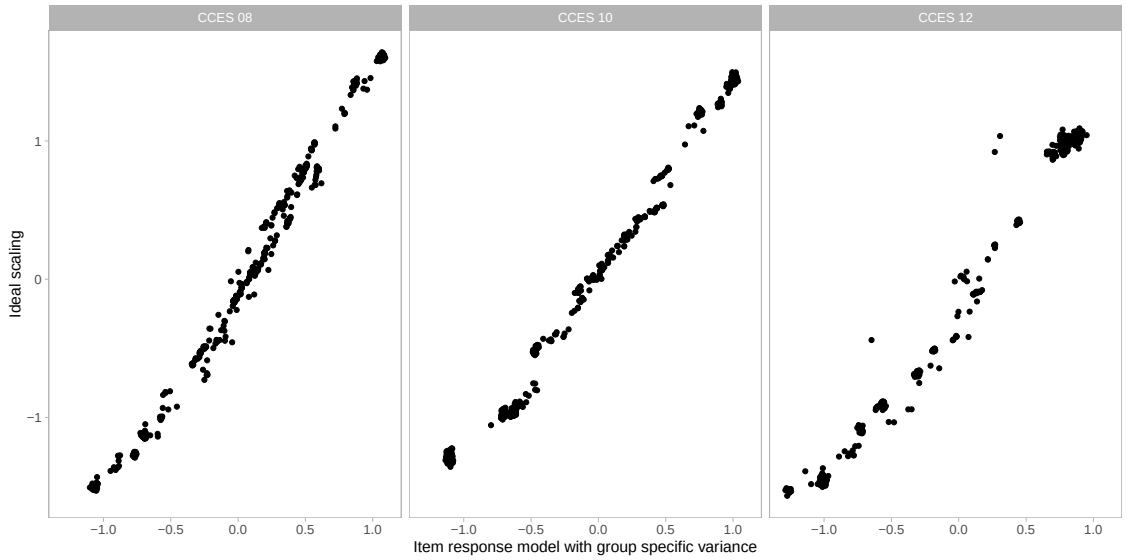

Figure 10: Latent trait estimates obtained from item response model with group specific variance and ideal scaling

E. Results from IRT with group specific variance depend the number of respondents included

I follow Jessee [forthcoming] and test if the estimation becomes "robust to seemingly innocuous factors such as the number of respondent included in the data"[Jessee, forthcoming, p.2]. The two group item response model is re-estimated for an increasing number of respondents, starting with half the number of respondents compared to legislators, same number of respondents, double and four times the respondents. If the method has potential for joint scaling the posteriori distribution of the estimated error variances should be relatively unaffected by this increase.

The results are reported in figure 11 and accentuate that the estimation is not robust to the number of respondents included. For all studies, the estimation identifies a decreasing difference between the error variances for the two groups. For the CCES 2010 this assimilation is the strongest. Allready when including twice as many respondents in the analysis, the error variance of the two groups are alike. For the CCES 2012 and

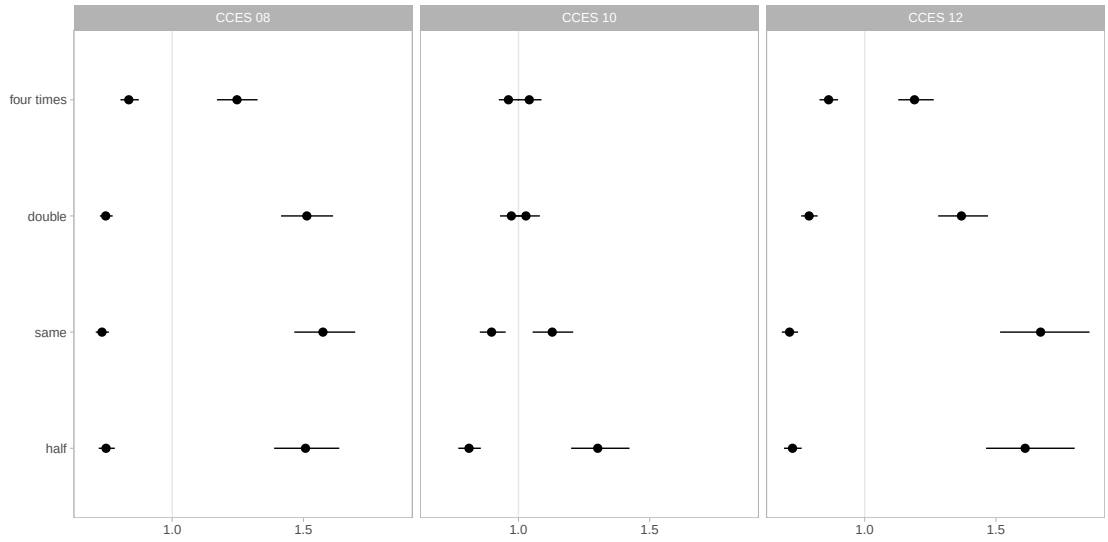

Figure 11: Parameter estimates for group specific error variance with varying number of respondents.

CCES 2008 they get far closer together with four times the respondents. Presumably, the gap would close when further increasing the number. In retrospect, these results are not surprising, as with an increasing number of respondents, data from the CCES dominates the estimation of the item parameters. The estimates then are similar to the once obtained for the respondents alone. The item parameters are less powerful in predicting votes for the legislators making them more similar to the citizens group. Because the error variances are re-transformed around one, the resulting estimation shows no difference anymore.