Towards a General Methodology of Bridging Ideological Spaces

Tzu-Ping Liu*[1], Gento Kato†[2], and Samuel Fuller‡[1]

[1]University of California, Davis
[2]Nazarbayev University

July 25, 2020

Abstract

Bridging ideal points estimated in two different groups of individuals (e.g., politicians and voters) into a common ideological space is an important, but relatively troubled branch of the ideological scaling literature. The most common procedure is *joint-scaling*, which is to merge two groups prior to the joint estimation of ideal points. This procedure has been criticized to ignore the possible differences in structures of ideological space between groups. Alternative approach, *dimensional-mapping*, accommodates this issue by estimating a transformation model of two ideological spaces. Existing dimensional mapping techniques, however, are of limited use, since they are not applicable to multi-dimensional and non-parametrically estimated ideal points and require "real" anchors (i.e., individuals who exist in two groups at the same time). Therefore, we propose a more generally applicable technique of dimensional-mapping that incorporates ideal points that deal with the above issues. Also, our method allows to subsample small numbers of individuals and set them as synthetic anchors: it does not require "real" anchors. We demonstrate the utility of our methodology by comparing its performance with that of joint scaling and existing dimensional mapping using two sets of voter-politician data from the United States and Japan. The result suggests that, while it makes less stringent assumptions than joint scaling and is more widely applicable than existing dimensional mapping techniques, our method can generate bridged ideal point estimates comparable to those generated from joint scaling.

*Ph.D. candidate, Department of Political Science. E-mail: tpliu@ucdavis.edu.

†Incoming assistant professor, Political Science and International Relations Department. E-mail: gento.badger@gmail.com.

‡Ph.D. candidate, Department of Political Science. E-mail: sjfuller@ucdavis.edu.

Introduction

Ideal point estimation is one of the most commonly used methods for ideological analysis in the field of political science. Directly derived from the spatial model of voting, it provides a theoretically and methodologically rigorous way to capture the underlying ideological preferences of different political actors (Poole 2005; Clinton 2012). In general, ideal point estimation methods determine the "ideological position" of an individual by exploiting large sets of either issue-specific preferences or political behavior across a sample of individuals (e.g. a chamber of Congress). These datasets range from roll-call votes in a parliament, opinions of judges in court cases, or responses to policy-related questions and feeling thermometers in public opinion polls. Assuming that the latent ideological space actually exists and matches the model's pre-specified structure (e.g., one versus two dimensions, different bases, or quadratic versus Gaussian utility functions), ideal points can be estimated based on the relative distances between every individual. These distances are derived from the relative and aggregate patterns in individuals' responses or behaviors, with those responding or behaving in similar fashions being grouped close together and those that are more different being placed further apart.

While conventional methods often make specific parametric assumptions regarding the structure of the estimated latent dimensions, such as the utility function for all individuals and the stochastic disturbances (random error) in preferences or behaviors (Poole and Rosenthal 1997; Heckman and Snyder 1997; Clinton, Jackman, and Rivers 2004), non-parametric methods relax many of these assumptions. These less stringent assumptions allows for a more appropriate and accurate estimation of ideal points among groups that likely have a weaker ideological structure than political elites, like voters (Poole 2000; Hare, Liu, and Lupton 2018).

A critical issue in ideal point estimation is the methods' dependency on the data on which they are applied. Given that ideal points are (most often) generated using relative Euclidean distances of the individuals in the given group/sample (e.g., voters or politicians), any two sets of ideal points estimated on *separate* groups (or even subsamples of those groups) are incomparable in most cases. This problem has been widely discussed in the literature, especially in the context of comparing the recovered ideal points of voters versus politicians. Current solutions to this incomparability problem can be divided into two general categories: **joint-scaling** and **dimensional-mapping**. The former addresses this issue by simply merging the two separate groups into one "pooled" data set prior to ideal point estimation (Bafumi and Herron 2010; Malhotra and Jessee 2014; Jessee and Malhotra 2013) while the latter preserves the *separate* estimation of ideal points and uses some sort of transformation method (e.g., an ordinary least squares (OLS) regression) to "bridge" the two ideological spaces.

Importantly, dimensional-mapping can further be divided into two separate, distinct approaches: **space-prediction** and **linear-mapping**. The first approach, space-prediction, estimates ideal points on only one group (e.g., politicians) and then utilizes these estimated (posterior distributions of) parameters to predict the ideal points of the other group (e.g., voters) (Jessee 2016). The second approach, linear-mapping, requires "anchors" (i.e., common individuals in both groups): After separate estimations of the two groups' ideal points, a transformation model is estimated based on anchors' positions in two spaces (e.g., Shor, Berry, and McCarty 2010).

While relatively effective in some situations, all of these approaches suffer from significant limitations. To start, the key assumption of joint-scaling is that the "ideological space underlying the preferences of the two groups is structured in the same way" (Jessee 2016, 1110). If this assumption is violated, ideal points estimated through joint-scaling are sensitive to trivial factors such as the relative size of the two groups. Dimensional-mapping does not suffer from this issue, but this method is of limited use when ideal points are multi-dimensional and/or should be estimated non-parametrically. Space-prediction is simply not a possibility in these scenarios since non-parametric estimations of ideal points do not generate parameters for predicting the second group's ideal points. Linear-mapping is a possibility here, but previous methods have only been applied in one-dimension with no explicit generalization developed for multi-dimensional ideal points.

In this paper, we introduce a new linear-mapping method that accommodates non-parametric *and* multi-dimensional ideal point estimation to reduce the limitations present for current bridging methods. In addition to addressing these issues, our method only requires two groups (e.g., politicians and voters) to have a common set of votes/questions and does not necessarily require pre-defined anchors. If pre-defined anchors do not exist, our method randomly selects a small number of individuals from one group (e.g., voters), merges them with the other group (e.g., politicians), and treats them as synthetic anchors. Then, after estimating ideal points for each group, we apply Procrustes transformation to estimate the transformation matrix based on the anchors' coordinates in two spaces. This matrix is then used to transform the rest of ideal points from the space of one group (e.g., voters) to that of another group (e.g., politicians).

This article proceeds as follows. First, we review existing bridging methods and describe in detail their limitations. Next, we introduce our new methodology. We then demonstrate its performance through simulated data experiments. In addition, we apply our method to two datasets that allow for bridging between politicians and voters: the 2004-2005 Senate Representation Survey in the United States (Jessee 2009) and 2012 UTokyo-Asahi Survey (UTAS) in Japan (Imai, Lo, and Olmsted 2016). In both applications, two-dimensional ideal points are estimated through non-parametric (ordered) optimal classification (OC) (Poole 2000; Hare, Liu, and Lupton 2018) and the performance of our method is compared with those of joint-scaling and linear-mapping. Lastly, we conduct two Monte-Carlo experiments for evaluating how well our approach can recover the latent common space for different sets of ideal points, and for assessing to what extent the inclusion of extra samples impacts the original distribution of ideal points. We conclude by discussing both the advantages and limitations of our approach and its implications for future development of ideal point estimation techniques.

Existing Bridging Methodologies and Their Limitations

To overcome the data-driven nature of ideal point estimation, there exist two general types of bridging methods: *joint-scaling* and *dimensional-mapping*. Both methodologies share the requirement that either there exist common questions/votes between both groups and/or that there are common individuals shared across both groups. The primary difference between the two methods is how the two groups are merged or bridged. Joint-scaling combines data from two (or more) groups with the same or similar set of either roll-call votes or survey questions into one *joint* dataset and then uses standard methods to estimate the combined ideal points. This method is by far the simplest: one merely need merge the two datasets and then they could apply any ideal point estimation method. On the other hand, joint-scaling makes the strong assumption that the ideological space of each combined group has an identical structure, i.e., the two sets of ideal points are generated from the same data generating process (DGP) (Jessee 2016). If this assumption is not satisfied, which is likely for the case when we compare highly ideologically constrained politicians versus less constrained voters, the structure of the jointly estimated ideological space can only *partially reflect* the true structure of ideological space for any one group. Additionally, since a jointly estimated ideological space is influenced by the relative sizes of each group the resulting structure can be highly skewed towards the larger group if they do not share the same underlying structure (e.g., a sample of 500 voters can overwhelm a sample of 100 senators).

In contrast, *dimensional-mapping* methods first estimate separate ideological spaces for each group and then transforms one group ideal points onto the other group's space. This procedure can thus accommodate the potential differences in the structures of groups' different ideological spaces. Within dimensional mapping, there exist two popular approaches: *space-prediction* and *linear-mapping*. Space-prediction uses the model parameters estimated on one group to bridge the two groups by simply plugging in the other group's data to generate their ideal points in the same space. For example, to bridge voters and politicians, Jessee (2016) trains a Bayesian probit-link ideal point model (Clinton, Jackman, and Rivers

2004) only on voters (politicians) and then uses the estimated parameters in each iteration to predict the ideal points of politicians (voters). While intuitive, there are limitations to this approach. One major issue is that it (almost) completely ignores the structure of ideological space for the *transformed* group. Since the core idea of bridging is to recognize the intrinsic differences between and within subspaces, ignoring one subspace makes this approach at best incomplete as a bridging method. Finally, space-prediction is simply unavailable for non-parametric ideal point methodologies, as these methods do not estimate model parameters.

Finally, linear-mapping estimates ideal points within both groups first and then estimates a transformation matrix to merge the two groups. Its biggest requirement is that there must exist "anchors," or common individuals who exist in both groups. After estimating the ideal points of the two groups separately, a transformation model is estimated based on anchors' coordinates in the two ideological subspaces. The transformation matrix is then used to map the rest of the ideal points onto one of the subspaces. This approach overcomes the limitations that plague joint-scaling and space-prediction as it uses the structural information from both groups' spaces to bridge between the two. Currently, however, only one technique exists for linear-mapping using a bivariate linear regression on a one-dimensional ideal point space (Shor, Berry, and McCarty 2010; Shor and McCarty 2011). This approach is thus limited by the use of parametric linear regression and it is not immediately clear whether this technique can be generalized to apply to multi-dimensional spaces. Another limitation for this approach is the pre-defined anchor requirement, which is rarely met, especially when comparing elite- versus mass-level ideological spaces (e.g., legislators and voters).

A General Methodology of Linear Mapping

In this section, we introduce our generalized linear-mapping methodology with an application to a two-dimensional (2-D) ideological space.[1] In a two-group bridging scenario, our method

[1]Our method can easily be generalized to higher dimensional spaces as well.

can be described as merging two separate 2-D subspaces with ideal points. One obvious question is how can we generate one combined space that recovers the *complete view* of ideal points? Overall, our method works in two stages to achieve this: First, we randomly generate overlapping data points, anchors, in each of the two groups and then estimate the ideal point in two separate 2-D subspaces. Second, we then use the overlapping data points to re-scale the subspaces to finally combine them. The following sections describe each stage in detail.

For the purposes of illustration, assume that we have two separate, unique sets of 1,000 survey respondents sharing an identical set of 30 questions. The first step is to take a few, say 20, respondents from the first survey and treat them as "synthetic" anchors by inserting them into the second survey's dataset. These synthetic anchors allow us to overcome the traditional linear-mapping requirement of shared respondents and thus merge the two spaces. We then apply non-parametric ideal point estimation, such as optimal classification (OC) or ordered optimal classification (OOC), depending on the data type, on each survey separately. This step not only non-parametrically estimates the coordinates (2-D ideal points) of respondents but also constructs two separate 2-D Cartesian coordinate systems with independent bases. In other words, this remedies the problems that plague joint-scaling and space-prediction in which either one space dominates the ideal point estimation of the other (joint-scaling) or most of the structural information is lost in merging the two groups (space-prediction). Simply put, these separate independent coordinate systems retain the structural information present in both groups while also enabling higher-dimensional ideal point estimation.

However, to use synthetic anchors we must make the assumption that the 20 first-survey respondents transferred into the second survey would not have different responses if they were indeed in the second survey. This assumption is analogous to the one required to use real, pre-defined anchors in bridging two groups. For example, Shor, Berry, and McCarty (2010) use legislators who served in both state-legislatures and Congress as anchors to connect state and congressional ideological spaces. Although these legislators are "real" anchors, the authors must still assume that these bridging legislators make similar decisions, using the same underlying ideological space, regardless of whether they are in the state-legislature or the US Congress. Thus, using real anchors does not exempt researchers from the assumption that anchors' ideological preferences are stable across time, group memberships, and/or institutions. Finally, given that ideal point estimation is data dependent, we must make the additional assumption that the ideological structure present in the second survey is not sensitive to the inclusion of synthetic anchors.[2]

In the next step, we utilize these 20 anchor-pairs to estimate a transformation matrix which transforms the estimated coordinates from one space onto the other. when transforming a 2-D space with **homogeneous coordinates**, the procedure can be explained, without a loss of generality, by the equation $\mathbf{d}_1 = \mathbf{T}\mathbf{d}_2$:

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix}$$

where $\mathbf{d}_1 = (x, y)$ and $\mathbf{d}_2 = (x', y')$ are a pair of corresponding points in the two subspaces separately, and the matrix with nine parameters, $\mathbf{T}$, is the transformation matrix. When the depth of each subspace is irrelevant, i.e., when conducting one- or two-dimensional ideal-point estimation, one usually leaves the three parameters on the last row of $\mathbf{T}$ as constant, usually $h_{31} = 0$, $h_{32} = 0$, and $h_{33} = 1$, and only estimate the rest six parameters, which is called (general) affine transformation.

Here, the existing methodology uses bivariate linear regression which only applies to one-dimensional scaling. However, if we instead apply multiple linear regression to estimate the

[2]However, as we show with our synthetic and real data applications, this assumption can be directly tested by comparing the estimated space that includes the anchors with the one that does not.

transformation parameters in a 2-D coordinate system, it can be defined as:

$$x = \alpha_{x0} + \beta_{x1}x' + \beta_{x2}y'$$
$$y = \alpha_{y0} + \beta_{y1}x' + \beta_{y2}y'$$

This formula shows that the use of linear regression in the transformation of a 2-D ideological space is equivalent to estimating the affine transformation matrix. When applied to the transformation of multi-dimensional spaces, affine transformation can, however, cause problems given that it allows shearing. Shearing (or shear mapping) uses a direction, or line, on which all points remain fixed, while adjusting all other points parallel to the direction by a distance proportional to the direction (see the Shearing panel in Figure 1 for an example of horizontal shearing) (Borg and Lingoes 2012; Solomon and Breckon 2011). Because shapes can geometrically describe the structure of a configuration (Solomon and Breckon 2011), which is estimated by the relative distances between data points, once the configuration is sheared, the proportional distances between points (the original shape) is lost in transformation. Simply put, linear regression mapping in multiple dimensions necessarily loses the original shape of the ideal point data through transformation.

To avoid this significant issue and preserve the original relationships between data points in the transformed group (i.e., the second survey in our illustration), we adopt Procrustes transformation to estimate the transformation matrix. Given two sets of correspondent point clouds, this method finds a transformation matrix which fits a configuration to the other as closely as possible using only scaling (changing the size of a shape without changing its centroid and structure), rotation (rotating a shape without changing its centroid and structure), translation (moving the centroid of a shape without changing its structure), and reflection (flipping a shape over a line without changing its size and structure), *not* shearing (Borg and Groenen 2005; Cox F. and Cox A. A. 2000).[3] Unlike affine or projective

[3]We present examples of each transformation in Figure 1.

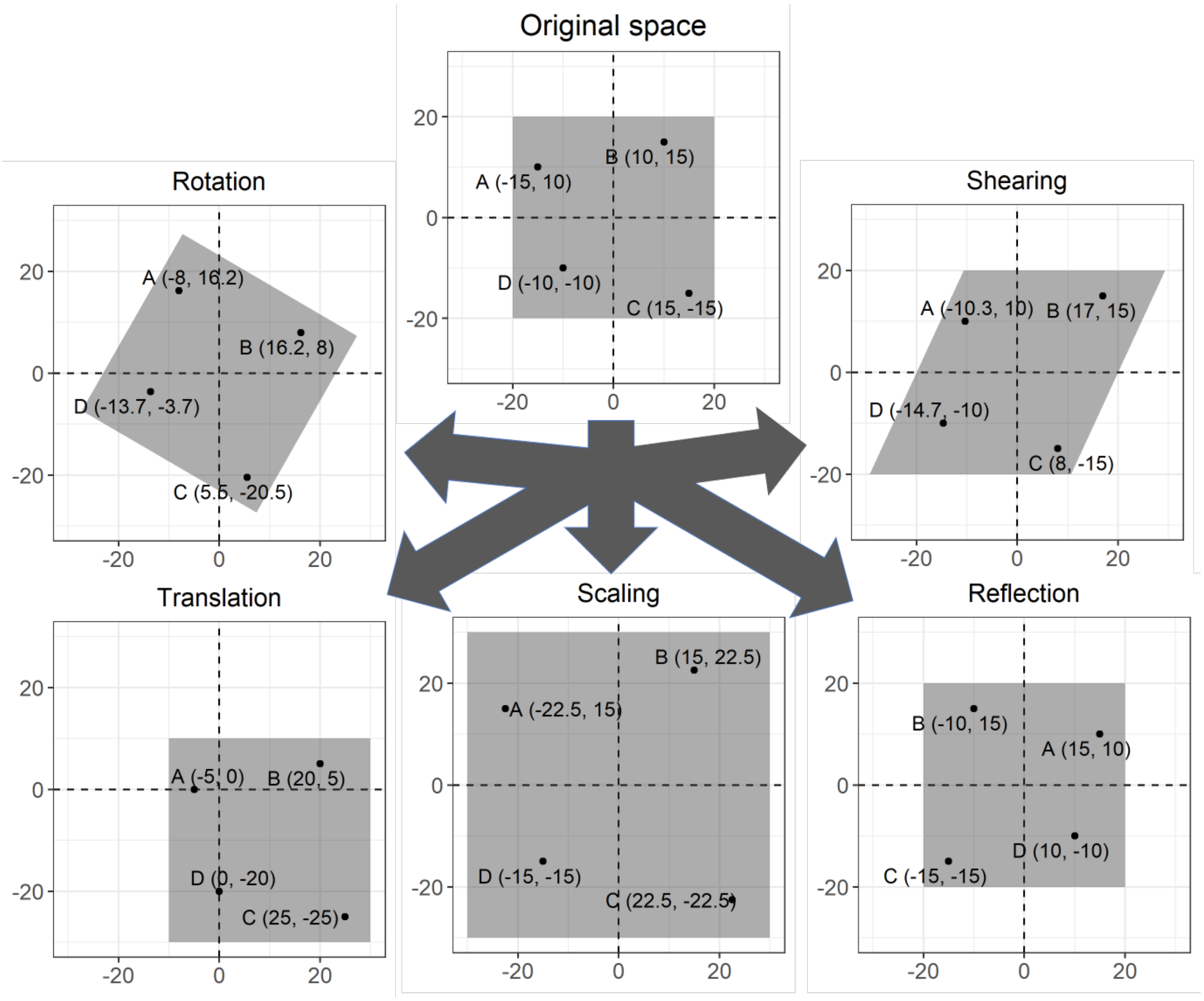

Figure 1: Geometric Transformations

transformation which allow shearing, Procrustes transformation preserves the original shape of the ideal point structure.

Formally, Procrustes transformation can be described as follows. Given two point clouds $\mathbf{X}$ and $\mathbf{Y}$, we would like to find a rotation matrix, $\mathbf{T}$, a translation matrix, $\mathbf{t}$, and a scaling scalar, s, such that $\mathbf{Y} \approx s\mathbf{X}\mathbf{T} + \mathbf{1}\mathbf{t}'$. Similar to ordinary least squares regression (OLS), we need to minimize the sum of squared errors to derive unknown parameters and matrices, but

Procrustes transformation does so through the Frobenius norm (Borg and Groenen 2005).[4]

$$\min_{||\mathbf{T}||^2=\mathbf{I}} \operatorname{tr}[\mathbf{Y} - (s\mathbf{XT} + \mathbf{1t}^{'})]^{'}[\mathbf{Y} - (s\mathbf{XT} + \mathbf{1t}^{'})] \tag{1}$$

The primary advantage of Procrustes transformation is that it maintains the projected points' original relationships in the new space. This is because Procrustes transformation does not change the shape of a given configuration, i.e., it preserves the proportion between edges of a configuration (the shape). Simply put, Procrustes transformation, above and beyond bivariate and multiple regressions, allows us to examine ideal points and their relative structure (shape) through the perspective of a (new) common space. The application of this transformation method is an important and novel contribution to the bridging literature in that it enables the non-parametric bridging of multi-dimensional spaces with no loss in the original shape of the respective ideological spaces.

Application

In this section we apply our method alongside previous approaches to demonstrate its utility across two different datasets, the 2004–2005 Senate Representation Survey and the 2012 Japanese UTAS Survey. Each alternative strategy generates OC or OOC ideal points for both voters and politicians in a common space based on a dataset with (1) voters and politicians pooled together, (2) only voters, and (3) only politicians. As described earlier, our methodology makes fewer assumptions than each of the alternative strategies when bridging two separate sets of actors into a common space.

Case 1: American Voters and Senators

We first examine the data also used in Jessee (2016) that surveys voters and politicians on an identical set of questions. This survey was conducted between December 2005 and January

[4]The restriction of $||\mathbf{T}||^2 = \mathbf{I}$ is necessary or this equation cannot be solved. Furthermore, this restriction guarantees that the derived rotation matrix, $\mathbf{T}$, is an orthogonal matrix.

2006, and includes questions corresponding to specific roll-call votes that also occurred in the U.S. Senate between 2004 and 2005. Combined with Senators' votes, this survey records voters' and senators' stances on each of these votes. Given that the survey was originally fielded to examine voters' perceptions of their own senators, the sample includes, at a minimum, 100 respondents from each state. We present the results of this analysis in Figure 2.

Figure 2 can be divided into three sets of analyses. We first present the results of joint scaling in the center panel. Second, we map Senators' ideal points to voters' latent ideal space by randomly selecting twenty voters to be used as anchors. The top row of the right panel presents the groups' separately estimated ideal points. The middle and bottom rows present the transformed ideal points onto the *voters*' latent space through Procrustes transformation and regression, respectively. Finally, we map voters' ideal points to senators' latent ideal space by randomly selecting twenty senators to be used as anchors, with the left panel presenting the groups' separately estimated ideal points. Again, the middle and bottom rows of the left panel present the Procrustes-transformed and regression-transformed voter ideal points on the *senators*' latent space, respectively.

The results of Figure 2 demonstrate how ideal point estimation is heavily data-driven. As one can see, respondents' estimated locations are not identical in these different combinations of data. In other words, the results indicate that when researchers utilize incorrect assumptions to estimate ideal points, these locations will inevitably be biased/distorted. Furthermore, as Jessee (2016) argues that when merging datasets for scaling we are necessarily making a compromise between the dimensional structure independently present in the two different datasets, the results show that relatively larger group (i.e., the voters) is distorted relatively less than the smaller group when estimating their ideal points.

Additionally, the bottom rows of the left and right panels in Figure 2 demonstrate the potential issue with using regression transformations in a multi-dimensional space. As discussed, shearing distorts the original shape of the ideal point structure, distorting important information from the original estimations. Although in the end ideal points of one of the

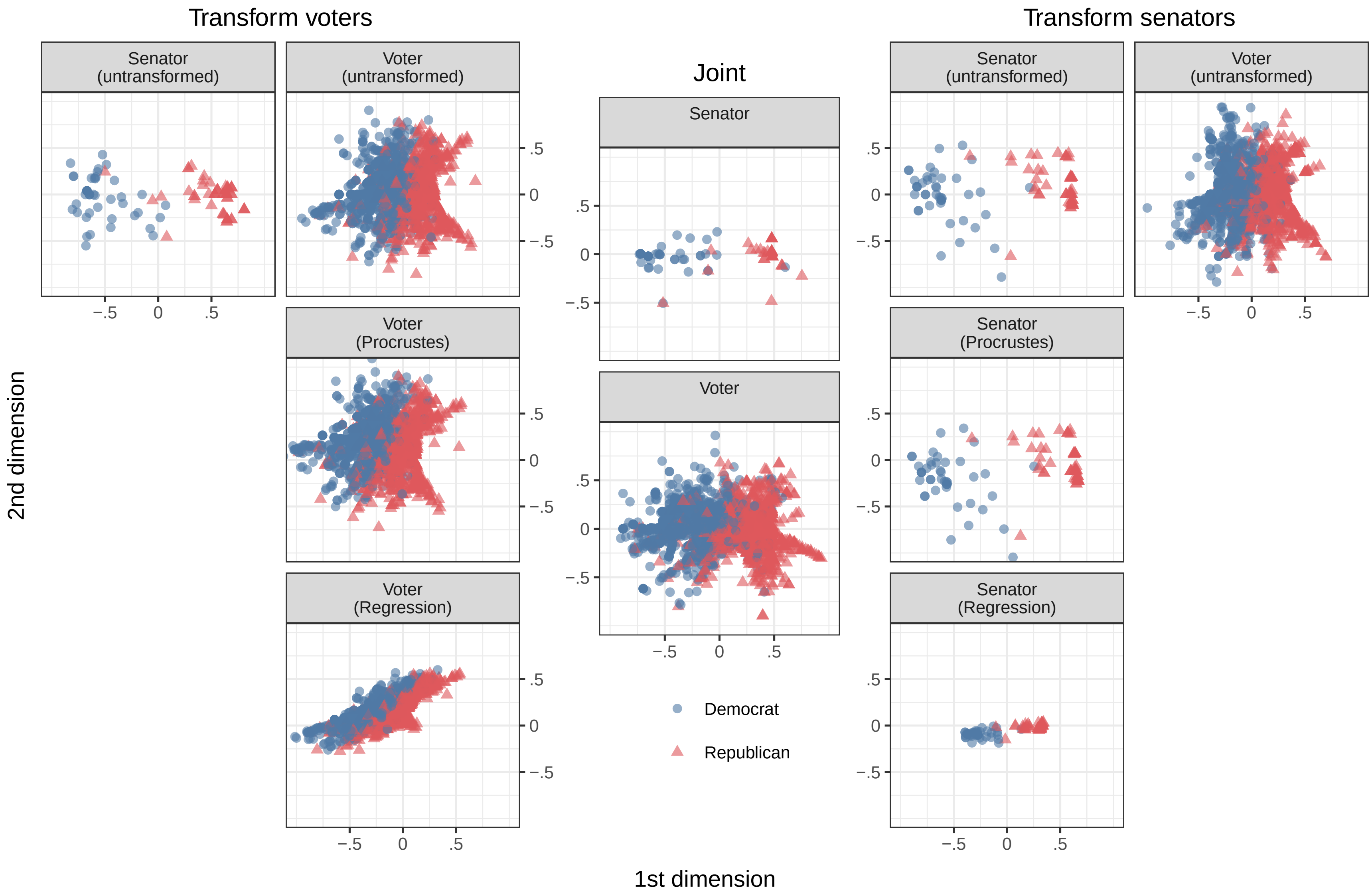

Figure 2: Senator Data OC Results

Table 1: Centroids by Partisan and Group (Jesse)

Partisan	Group	Combination	Coordinates	Distance
Republican	Senator	Senator to Voter	$(0.52, -0.03)$	1.09
Democrat	Senator	Senator to Voter	$(-0.55, -0.21)$	1.09
Republican	Voter	Senator to Voter	$(0.14, 0.00)$	0.29
Democrat	Voter	Senator to Voter	$(-0.16, 0.03)$	0.29
Republican	Senator	Voter to Senator	$(0.53, -0.03)$	1.06
Democrat	Senator	Voter to Senator	$(-0.53, -0.04)$	1.06
Republican	Voter	Voter to Senator	$(-0.04, 0.13)$	0.28
Democrat	Voter	Voter to Senator	$(-0.31, 0.21)$	0.28

groups are successfully projected onto the latent space constructed by the other group, the transformed figures are somewhat compressed making them relatively uninformative. In contrast, as claimed above, the middle rows of the left and right panels in Figure 2 illustrate that Procrustes transformation not only transforms ideal points of one latent space to the other, but also preserves the original configurations/information as much as possible.

Once two sets of ideal point clouds are transformed on the same latent space, we can further calculate the *centroid* of each group by partisanship. We can then use this information to determine the degree of polarization between the two groups by calculating the Euclidean distance between the groups' centroids. To do so, for instance, the centroid of Republican Senators' ideal points, C_{rs}, is calculated as $C_{rs} = (\bar{x_{rs}}, \bar{y_{rs}})$, where $\bar{x_{rs}}$ is the mean x-coordinate and $\bar{y_{rs}}$ is the mean y-coordinate of all Republican Senators. We present the calculated centroids and Euclidean distances between parties in Table 1.

Figure 2 and Table 1 together illustrate that, regardless of how ideal points are transformed, that even though in absolute terms voters are polarized, the level of polarization is much greater among Senators than ordinary voters. Finally, and importantly, these results are consistent with a large amount of prior research documenting both polarization and ideological constraint in the mass public versus elites (Lupton, Myers, and Thornton 2015; Hare et al. 2015). This provides further qualitative evidence that our results are working well in bridging between two different groups while maintaining their structural differences.

Case 2: Japanese Voters and Candidates

In this section we analyze the UTokyo-Asahi Survey (UTAS), fielded in Japan during the House of Representatives election in 2012.[5] At each election, UTAS contains two sets of surveys—voters and candidates—that share an identical set of policy questions. All policy questions are answered on an ordinal scale and are thus compatible with OOC. The voter survey randomly samples 3000 respondents from Japan's list of registered voters' and responses are filled out using a mail-in questionnaire. The response rates are 63.3% (N=1900) in 2012. The candidate survey is sent to all congressional candidates in a given election. The response rates are 93.4% (N=1404) in 2012.[6]

Somewhat obviously, we focus on bridging between voters and politicians in this Japanese context. Bridging voters' and politicians' ideologies is a particularly interesting question in Japan, for at least two reasons. First, in contrast to the United States, parties' positions on the ideological spectrum are not always obvious or clear to voters in Japan (Jou and Endo 2016). Second, a large portion of Japanese voters are considered to be "independent" voters (called *muto-ha*), and Japanese politicians are also known to frequently switch (or form new) parties. Combined, these facts indicate that in Japan, it is very difficult to assume that voters and politicians share a common ideological space. Our method, when compared to others, is much more suited for these situations and thus should be more valid.

Similar to Figure 2, again, Figure 3 demonstrates that the derived ideal points are highly data-driven, i.e., ideal points are strongly biased when using incorrect assumptions. Furthermore, the issue with regression transformation persists—the transformed configuration is too compressed to be informative. As shown in bottom panels of the figure, regression transformation tends to project ideal points almost solely onto the x-axis and distorts the information conveyed originally by the y-axis.

[5] UTAS is conducted by Masaki Taniguchi of the Graduate Schools for Law and Politics, the University of Tokyo and the Asahi Shimbun. The original dataset is available from the survey's website (`http://www.masaki.j.u-tokyo.ac.jp/utas/utasindex.html`).

[6] We also analyzed UTAS 2009 and found very similar results. For the sake of brevity, we present these results in online appendix.

Table 2: Centroids by Partisan and Group (UTAS12)

Duo	Partisan	Group	Combination	Coordinates	Distance
LDP v.s. DPJ	LDP	Candidate	Candidate to Voter	$(0.17, 0.39)$	0.51
	DPJ	Candidate	Candidate to Voter	$(-0.17, 0.01)$	0.51
	LDP	Voter	Candidate to Voter	$(0.07, 0.04)$	0.16
	DPJ	Voter	Candidate to Voter	$(-0.07, -0.02)$	0.16
LDP v.s. JRP	LDP	Candidate	Candidate to Voter	$(0.17, 0.39)$	0.60
	JRP	Candidate	Candidate to Voter	$(0.26, -0.20)$	0.60
	LDP	Voter	Candidate to Voter	$(0.07, 0.04)$	0.13
	JRP	Voter	Candidate to Voter	$(0.10, -0.08)$	0.13
DPJ v.s. JRP	DPJ	Candidate	Candidate to Voter	$(-0.17, 0.01)$	0.48
	JRP	Candidate	Candidate to Voter	$(0.26, -0.20)$	0.48
	DPJ	Voter	Candidate to Voter	$(-0.07, -0.02)$	0.19
	JRP	Voter	Candidate to Voter	$(0.10, -0.08)$	0.19
LDP v.s. DPJ	LDP	Candidate	Voter to Candidate	$(0.23, 0.22)$	0.38
	DPJ	Candidate	Voter to Candidate	$(-0.00, -0.08)$	0.38
	LDP	Voter	Voter to Candidate	$(0.11, 0.09)$	0.18
	DPJ	Voter	Voter to Candidate	$(-0.04, -0.01)$	0.18
LDP v.s. JRP	LDP	Candidate	Voter to Candidate	$(0.23, 0.22)$	0.47
	JRP	Candidate	Voter to Candidate	$(0.28, -0.25)$	0.47
	LDP	Voter	Voter to Candidate	$(0.11, 0.09)$	0.14
	JRP	Voter	Voter to Candidate	$(0.14, -0.05)$	0.14
DPJ v.s. JRP	DPJ	Candidate	Candidate to Voter	$(-0.00, -0.08)$	0.32
	JRP	Candidate	Candidate to Voter	$(0.28, -0.25)$	0.32
	DPJ	Voter	Candidate to Voter	$(-0.04, -0.01)$	0.19
	JRP	Voter	Candidate to Voter	$(0.14, -0.05)$	0.19

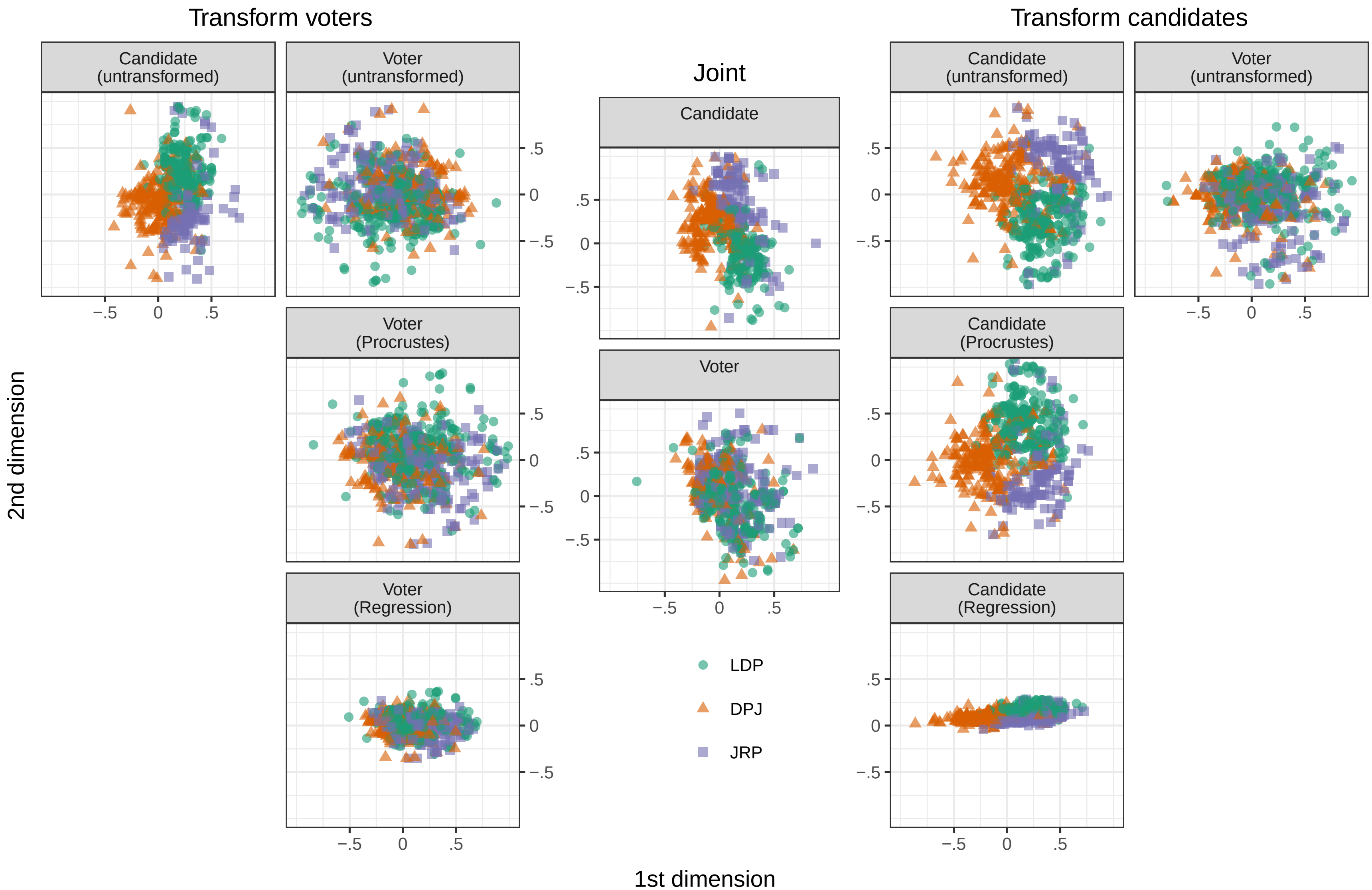

Figure 3: UTAS12 Data OOC Results

Combined Figure 3 and Table 2 illustrate the usefulness of our methodology above and beyond standard methods when two bridged groups may not share the same underlying ideological space. Furthermore, we conclude that first, in Japan, party lines do accurately reflect their policy outcomes for candidates/politicians, i.e., political polarization does exist among Japanese politicians. Nevertheless, compared with the 2005–2006 Senate Representation Survey, the results demonstrate that the level of polarization among Japanese elites is much smaller than those in the U.S. Furthermore, the derived results indicate that Japanese voters are graphically and geometrically overlapping on this latent ideological space. This is reflected in the small distances between the ideal point centroids of different parties' voters, further illustrating that party lines do not accurately divide voters' policy preferences.

This is a perfect example, in contrast to the American case, in which two or more dimensions may be needed to understand latent dimensional spaces. Whereas in the American situation the first dimension cleanly splits partisans, the Japanese case is a clear example of when a single dimension is uninformative. Overall, while not only providing evidence of our method's utility and robustness, the Japanese case further supports the importance of being able to bridge multi-dimensional spaces. As our experiments in the next seciton will also show, our method makes an important contribution in enabling multidimensional linear-mapping. In conclusion, unlike American voters, Japanese voters' political preferences have not been clearly polarized or divided along party lines and, overall, they generally hold moderate views across all policies.

Validating the Use of Synthetic Anchors

In our applications, we use synthetic anchors, as oppose to *real* anchors, to bridge ideological spaces. This procedure makes an assumption that the inclusion of out-group extra samples do not distort the estimation of original latent space. In this section, we test this assumption in UTAS 2012 dataset, following the approach taken in Jessee (2016). Specifically, we estimate voters' two-dimensional ideal points (N=1900) through OOC by including zero, twenty, fifty,

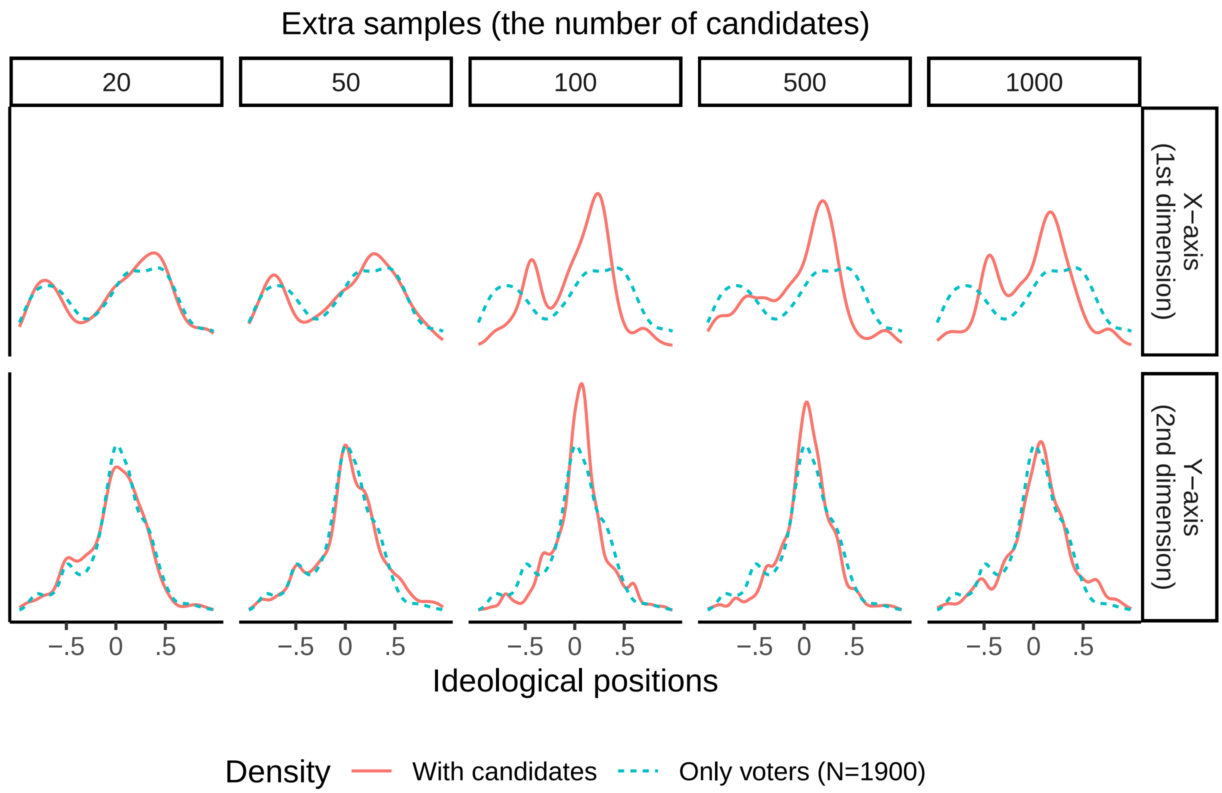

Figure 4: Density Plots of Voters' Positions with Different Amount of Extra Samples from UTAS 2012

one-hundred, five-hundred, and one-thousand randomly sampled candidates. We repeat this procedure four-hundred times[7] and assess if distributions of voter's ideologies changes by the inclusion of extra candidate samples.

Results are presented in Figure 4. In each panel, a solid line shows the density distribution of voters' ideal points estimated with the inclusion of extra candidate samples, averaged across four-hundred experiments. A dashed line indicate the density distribution of voters' ideal points without any candidate samples. Given that we estimated two-dimensional ideological space, the top row shows distributions for first dimension (i.e., x-axis), the bottom row for second dimension (i.e., y-axis). Columns indicate the number of extra candidate samples included in estimation.

[7]Given that there is no random sampling involved, voters' ideal points without candidate samples are only estimated once.

We see that in general, the larger the candidate samples added to voter data, the more distorted the original distribution of voters' ideologies (i.e., distribution estimated without candidate samples). For derived coordinates on the x-axis shown in the top row, the density is not significantly different than the original when the sample-size is smaller than fifty. As the number of extra samples grows, however, the two peaks of the original bi-modal density gradually converge around the origin (0,0). The distortion caused by synthetic anchors is less severer on the y-axis (the bottom row). Still, density distributions with extra samples tend to deviate more from original as the size of the candidate sample grows.

The results of the above Monte-Carlo experiment provide an important rule of thumb regarding the size of the anchor sample when using our method. As the results show, as long as the size of the synthetic anchor sample is under a reasonable level (approximately under 3.5% or fifty in our sample of 1,400), the original density of estimated ideal points is not significantly distorted (see the bottom row of Figure 4). In fact, the results demonstrate that even when there are no real anchors, synthetic anchors enable us to accurately bridge two separate datasets and map them onto the same latent space.

Compare Performances with Simulated Data

While we have demonstrated the utility of our approach, we have yet to address whether this general approach performs well to reveal a true latent space underlying individuals' preferences. In this section, we design an experiment to address this concern.

To verify that the general approach can accurately reveal a real latent space, we adopt an informal Monte-Carlo experiment similar to Hare, Liu, and Lupton's (2018).[8] The simulation uses 2,400 "respondents" and a set of 40 five-point Likert scale policy questions (like a standard survey). We also, separately, generate each respondent's ideal point coordinates in a two-dimensional space by drawing from a bivariate normal distribution centered at zero with correlations between the two dimensions also drawn from a uniform distribution

[8] See Hare, Liu, and Lupton's (2018) for a detailed explanation for this experiment.

between -0.1–0.7. The "policy question normal vectors" are generated by drawing from the edges of a unit circle in 2-D space, then each of the five possible responses to each issue are randomly selected and projected onto their resepective normal vectors.

To generate each respondent's response to each question, we first generate each response level's location (one to five) along with each normal vector based on the distribution of all respondents' locations along that normal vector. We then randomly assign one of three distributions, the quadratic, normal, and linear distribution (with equal probability and the same mean and similar variances) to each respondent, and further generate each respondent's response based on the probability corresponding to the distance between each level and that respondent's location.[9] For instance, given a respondent's distances between each of five levels, we can derive the probability of each level for that respondent based on his/her designated distribution, and further randomly pick one level as their actual choice based on each level's corresponding probability.

After the data is generated, we randomly split the data into two subsets, one containing 1,400 respondents and the other 1,000. We employ multiple bridging methods to estimate and merge ideal points on these two datasets, including:

1. Joint Scaling: Merge the two datasets and jointly estimate ideal points.

2. Procrustes Transformation: Our dimensional-mapping approach described above (with twenty extra anchors).

3. Linear Regression: Existing dimensional-mapping approach described above (with twenty extra anchors).

4. Concatenation: Estimate ideal points separately on two datasets and then concatenate them.

[9] Although the three utility functions are substantively different after the Gaussian's inflection points, OC or OOC constrains the boundary of ideal-points to a unit-circle. Under this unit-circle condition, these three utility functions should be exceedingly similar due to the ideal-points' close proximity to the mean.

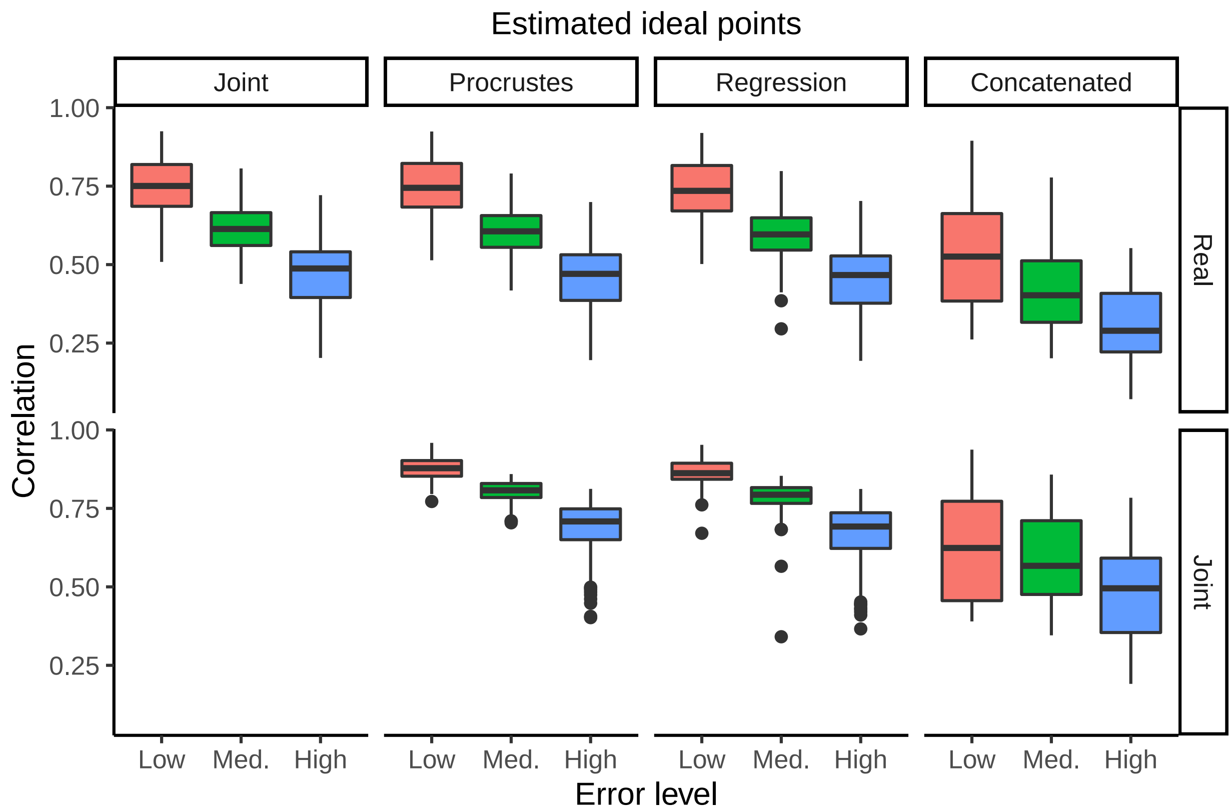

Figure 5: Correlation between Estimated and Real Positions

Each approach is repeated four hundred times. After all the results are derived, we further calculate correlations between the distance matrices between estimated ideal points from each approach and the real ideal points. The results of this experiment are presented in Figure 5.

Based on the proportion of incorrect choices (errors), we group results into three categories: low, medium, and high error. We then calculate two types of correlations by each error group. The first includes correlations between the real ideal points and the distance matrix of each method. The second compares the utility of the different methods through the correlations between the distance matrix of joint scaling and the distance matrices of the other methods.

As the boxplots in the top panels of Figure 5 demonstrate, first of all, joint scaling, Procrustes transformation, and regression transformation all perform similarly across the different error levels. Given that all respondents' ideal points are generated from similar distributions, the results of joint scaling also show the utility of OOC as demonstrated in Hare, Liu, and Lupton (2018). Furthermore, the results from Procrustes and regression transformation both indicate that with appropriate procedures, two different sets of ideal points can still be projected onto the same latent space.

The results from the separate estimations further validate that the results of the two transformation approaches are not simply due to the fact that the two sets of ideal points are generated from the same distribution. As discussed, given that the ideal points are estimated through the relative distances between individuals, the estimated results are highly data driven, i.e., different combinations of individuals will generate different latent spaces. The poor performance of the separate estimations illustrates that even though two sets of ideal points are generated from the same distribution, the estimated ideal points actually exist in different latent spaces due the structure of data. Furthermore, the correlations between the results of joint scaling the the rest of approaches indicate that joint scaling and the two transformation approaches all derive similar configurations of estimated ideal points.

By and large, the results of Figure 5 provide one important insight into bridging methodologies—that is, even when two groups meet the assumptions required by joint scaling, applying bridging methods will generate similar results, but if these assumptions are not met, bridging methods will yield vastly different and better results. Simply put, when trying to compare two sets of ideal points without knowing the underlying data generating process (which is most often the case), one should always use an appropriate bridging method, rather than joint scaling, to guarantee valid estimates and results.

Overall, the results of current Monte-Carlo experiment addresses important concerns regarding our method. When the two sets of ideal points are generated by the same data generating process, our results demonstrate that dimensional-mapping performs no worse than joint scaling. Furthermore, the results also show that the estimation of ideal points is highly driven by the combination of respondents, regardless of their distribution.

Discussion: Generalizing a Methodology of Bridging

In this paper, we develop a generalized approach to ideal-point bridging to address major issues and limitations present in previous methods. Compared with existing approaches, this generalized approach has four major advantages: 1) it can employ synthetic anchors when no real ones exist between the bridged groups; 2) this method can be applied to multi-dimensional latent spaces, which is often necessary for recovering meaningful dimensions (e.g., Japan); and 3) it preserves important information regarding the separate ideological structure of the different groups as much as possible.

The benefits of this method are also fueled in large part due to the significant threat of bias arising from the highly data-driven structure of ideal-point estimation. Specifically, our applications and experiments demonstrate that: 1) when two sets of data are not generated by the same distribution, joint scaling should not be used due to the bias caused by the data-dependent nature of ideal-point estimation; and 2) when two sets of data are generated by the same data generating process, dimensional-mapping performs just as well as joint-scaling. In fact, our method improves upon joint-scaling's performance when the data generating processes are heterogeneous and matches it when the DGPs are homogeneous. Given that prior knowledge of groups' DGPs is rare, at best, our findings suggest that one should almost always use linear mapping, generally, and our method, specifically, as demonstrated in Figure 5. This ensures a high degree of performance and the lowest possible risk of biased estimation.

Finally, although this method is primarily designed for non-parametric ideal-point estimation, we believe that this general approach can also be applied to parametric methods. Theoretically, because parametric methods assign specific DGP to the distribution of data, as compared with non-parametric methods, the post-hoc density of ideal points should be more stable if applied to parametric methods (i.e., the original density is not significantly distorted when using a reasonable amount of synthetic anchors). In other words, synthetic anchors should not significantly impact original data's structure, thus allowing researchers to apply this general approach to other parametric methods of ideal point estimation, such as the Bayesian item response model (Clinton 2012) or the blackbox scaling in basic space (Poole 1998).

References

Bafumi, Joseph, and Michael Herron. 2010. "Leapfrog Representation and Extremism: A Study of American Voters and Their Members in Congress." *American Political Science Review* 104 (3): 519–542.

Borg, Ingwer, and Patrick JF Groenen. 2005. *Modern Multidimensional Scaling: Theory and Applications.* Springer Science & Business Media.

Borg, Ingwer, and James Lingoes. 2012. *Multidimensional Similarity Structure Analysis.* Springer Science & Business Media.

Clinton, Joshua D. 2012. "Using Roll Call Estimates to Test Models of Politics." *Annual Review of Political Science* 15 (1): 79–99. eprint: https://doi.org/10.1146/annurev-polisci-043010-095836.

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

Cox F., Trevor, and Michael Cox A. A. 2000. *Multidimensional Scaling.* CRC Press.

Hare, Christopher, David A Armstrong, Ryan Bakker, Royce Carroll, and Keith T Poole. 2015. "Using Bayesian Aldrich-McKelvey Scaling to Study Citizens' Ideological Preferences and Perceptions." *American Journal of Political Science* 59 (3): 759–774.

Hare, Christopher, Tzu-Ping Liu, and Robert N Lupton. 2018. "What Ordered Optimal Classification Reveals about Ideological Structure, Cleavages, and Polarization in the American Mass Public." *Public Choice* 176 (1-2): 57–78.

Heckman, James J., and James M. Snyder. 1997. "Linear Probability Models of the Demand for Attributes with an Empirical Application to Estimating the Preferences of Legislators." *The RAND Journal of Economics* 28:S142–S189. doi:10.2307/3087459.

Imai, Kosuke, James Lo, and Jonathan Olmsted. 2016. "Fast Estimation of Ideal Points with Massive Data." *American Political Science Review* 110 (4): 631–656.

Jessee, Stephen. 2009. "Spatial Voting in the 2004 Presidential Election." *American Political Science Review* 103 (1): 59–81.

———. 2016. "(How) Can We Estimate the Ideology of Citizens and Political Elites on the Same Scale?" *American Journal of Political Science* 60, no. 4 (October): 1108–1124.

Jessee, Stephen, and Neil Malhotra. 2013. "Public (Mis)Perceptions of Supreme Court Ideology: A Method for Directly Comparing Citizens and Justices." *Public Opinion Quarterly* 77 (2): 619–634.

Jou, Willy, and Masahisa Endo. 2016. *Generational Gap in Japanese Politics: A Longitudinal Study of Political Attitudes and Behaviour.* New York: Palgrave Macmillan.

Lupton, Robert N, William M Myers, and Judd R Thornton. 2015. "Political Sophistication and the Dimensionality of Elite and Mass Attitudes, 1980–2004." *The Journal of Politics* 77 (2): 368–380.

Malhotra, Neil, and Stephen A Jessee. 2014. “Ideological Proximity and Support for the Supreme Court.” *Political Behavior* 36 (4): 817–846.

Poole, Keith T. 1998. “Recovering a Basic Space from a Set of Issue Scales.” *American Journal of Political Science* 42 (3): 954–993.

———. 2000. “Nonparametric Unfolding of Binary Choice Data.” *Political Analysis* 8 (3): 211–237.

———. 2005. *Spatial Models of Parliamentary Voting.* New York, NY: Cambridge University Press.

Poole, Keith T., and Howard Rosenthal. 1997. *Congress: A Political-Economic History of Roll Call Voting.* OCLC: 318375105. Oxford [England]; New York: Oxford University Press.

Shor, Boris, Christopher Berry, and Noran McCarty. 2010. “A Bridge to Somewhere: Mapping State and Congressional Ideology on a Cross-Institutional Common Space.” *Legislative Studies Quarterly* 35, no. 3 (August): 417–448.

Shor, Boris, and Nolan McCarty. 2011. “The Ideological Mapping of American Legislatures.” *American Political Science Review* 105 (3): 530–551.

Solomon, Chris, and Toby Breckon. 2011. *Fundamentals of Digital Image Processing: A Practical Approach with Examples in Matlab.* John Wiley & Sons.