Testing the Measurement Invariance of Nativism*

Yikai Zhao, *Texas A&M University*

Objective. Survey questions regarding immigration have been frequently used to compare how mass publics think about nativism, but it remains unclear whether the resulting nativism measures are comparable across countries. In this article, we empirically examine whether nativism achieves scalar invariance—an important measurement feature for valid instrumentation—across very different country contexts. *Method.* We examine measurement invariance of a nativism scale using survey data from 22 countries representing 16,096 individual respondents. To test for invariance, we utilize multigroup confirmatory factor analysis (MG-CFA). *Results.* While the nativism scale fails to achieve full scalar invariance, we are able to establish partial scalar invariance, indicating the nativism scale is valid across many national contexts. Though highly correlated, the means of the latent nativist variable also provide different rankings than the means of an additive scale. *Conclusions.* With a few limitations, nativist sentiments can be meaningfully measured across political contexts. Across 19 of the 22 countries, respondents answered survey items on nativism in ways that suggest their answers reflect the same underlying latent construct.

Since Donald Trump's surprising victory in the 2016 U.S. presidential election, there have been countless studies examining the psychological, sociological, and political factors responsible for his electoral success. Scholars have identified economic marginalization, authoritarianism, racism, sexism, and a host of other factors that help to explain Trump's support (Carmines, Ensley, and Wagner, 2016; Hooghe and Dassonneville, 2018; MacWilliams, 2016; Mutz, 2018; Newman, Shah, and Collingwood, 2018; Schaffner, MacWilliams, and Nteta, 2017; Sides, Tesler, and Vavreck, 2017). Issues related to immigration feature in several of these studies, as Trump's nativist appeals encompassed concerns about public safety, stagnant wages, job loss, government incompetence, and racism. According to several accounts, nativism, broadly defined, was the driving force behind Trump's 2016 election (Young, 2016).

While it might be tempting to treat Donald Trump as an outlier, nativism extends beyond the water's edge of U.S. politics. Globalization, immigration, and economic stagnation have fed right-wing populist movements throughout the world, including (but not limited to) Brexit and the UK Independence Party (UKIP) in the United Kingdom and the National Front in France (Davidov and Meuleman, 2012; González and Young, 2017; Hatton, 2016; Hellström and Hervik, 2014; Ivarsflaten, 2005). From a survey research standpoint, measuring nativist sentiments across countries presents a challenge (Johnson, 1998; King et al., 2004).[1] Can the same survey questions, or the same scale, be used cross-culturally to measure the same underlying concept? Do survey respondents share an understanding

*Direct correspondence to Yikai Zhao, Department of Communication, Texas A&M University, College Station, TX 77845 ⟨zdw2126062@tamu.edu;zdwzhaoyikai@gmail.com⟩. The author wants to extend special thanks to Dr. Kirby Goidel for his insights and expertise, which greatly assisted this research. The author is also grateful to Ipsos for sharing the data to make this analysis possible.

[1]King et al. (2004), for example, suggest using anchoring vignettes, an impossibility in the current study.

SOCIAL SCIENCE QUARTERLY

DOI: 10.1111/ssqu.12594

of survey questions? And, can we make scale comparisons across very different national contexts? In the parlance of psychometrics, we would describe a measure that is similarly understood and reported as invariant (or equivalent), meaning that the latent construct is the same (or nearly the same) across countries. Achieving measurement invariance (MI) gives us greater confidence that comparisons across groups, contexts, or time points are substantively meaningful.[2]

Questions related to MI have received a great deal of attention in psychology, education, and organizational behavior, but they have received less attention in political science (Alemán and Woods, 2016; Ariely and Davidov, 2011, 2012; Oberski, 2014; Oberski, Vermunt, and Moors, 2015; Reeskens and Hooghe, 2008).[3] This is unfortunate as researchers should not assume MI (or measurement equivalence) when examining items in cross-cultural contexts, particularly when they easily can construct tests for invariance using structural equation modeling (SEM) techniques (Meuleman, Davidov, and Billiet, 2009). When these tests indicate that measures are not invariant, observed differences across space and/or time may be the result of systematic biases, political contexts, cultural differences, language, or other factors related to respondents' understanding or interpretation of specific question wordings. When statistical tests indicate that MI has been achieved, researchers can have greater confidence that question items reflect the same underlying latent construct across contexts, time, and/or subpopulations. Establishing MI then is an important consideration when examining attitudes on issues or topics where comparability across countries, groups, or time is a key factor.

In the current study, our central question revolves around how individuals across countries interpret questions about nativism. Specifically, we test MI of a nativism scale proposed by Young (2016) and colleagues at Ipsos Public Affairs. Nativism, we should note, is not simply anti-immigration policy preferences or attitudes, but extends to a more generalized sense of threat targeted at immigrant groups or communities (Knoll, 2013; Mudde, 2007). Narrowly defined, anti-immigration sentiments tend to be a function of egocentrism, social distance, perception of ethnic threat, and avoidance of out-group contact (LeVine and Campbell, 1972). Nativism, in contrast, is rooted in fears of economic and cultural loss, especially the loss of cultural identity created by new waves of immigrants (Gratton, 2018; Inglehart and Norris, 2017; Mutz, 2018). We defer to other research regarding the *sources* of nativism; we are concerned here with whether this construct can achieve the level of invariance necessary to utilize it in analyses of public opinion cross-nationally. Thus, we ask two research questions:

RQ1: *Is the nativism scale equivalent across 22 countries? In statistical terms, does it achieve a level of scalar invariance that allows for meaningful comparisons across countries?*

RQ2: *To the extent that MI can be achieved, how does the level of nativism vary across country?*

Data and Methods

The data used in this analysis were collected by Ipsos Public Affairs and were based on surveys conducted from December 28, 2015 through April 16, 2016. Data were collected

[2]Our interest in the current article is measurement invariance across contexts. We note that it might also be used, for example, to test whether women and men or different racial groups respond similarly to a set of questions.

[3]In an exchange with Alemán and Wood (2016), Welzel and Inglehart (2016) argued that "constructs can entirely lack convergence at the individual level and nevertheless exhibit powerful and important linkages at the aggregate level."

TABLE 1

Correlations and Descriptive Statistics for Nativism Items

	M (*SD*)	Take Jobs	Take Services	Prioritize Hiring	Let in Immi-grants	Stop Immigration
Take jobs	2.9 (1.33)		0.73	0.59	0.16	0.66
Take services	3.1 (1.36)			0.58	0.21	0.68
Prioritize hiring	3.5 (1.32)				0.22	0.54
Let in immigrants	3.8 (1.18)					0.21
Stop immigrants	2.9 (1.34)					

online via the Ipsos online panel system and include 16,096 completed interviews with nonelderly adults from 22 countries. We used 19 of these countries in our subsequent analysis (South Korea, Japan, and Peru were excluded due to reliability concerns). In countries where Internet penetration rates are relatively high, the results should reflect the general population. In countries where Internet penetration rates are relatively low, the data likely reflect better educated and more urban populations rather than the general population. Final results were weighted to reflect recent population estimates for the respective countries.

Using items from the World Values Survey (WVS) and the General Social Survey (GSS), the Ipsos nativism scale captures perceptions that immigrants are taking jobs and social services away from native populations and that immigration is weakening the country. Specific question items are structured in agree/disagree Likert formats with respondents reacting to the following statements:

- Q1: *Take Jobs*: Immigrants take away jobs from real <Insert country>.
- Q2: *Take Services*: Immigrants take important social services from real <Insert country>.
- Q3: *Prioritize Hiring*: When jobs are scarce, employers should prioritize hiring people of this country.
- Q4: *Let in Immigrants*: <Country name> would be better off if we let in all immigrants who wanted to come. [Reverse coded.]
- Q5: *Stop Immigration*: <Country name> would be stronger if we stopped immigration.

For purposes of this analysis, we coded each individual item on a 1 to 5 scale with higher values indicating more nativist sentiments. "Don't know" responses, which reflect approximately 10 percent on each item, were recoded to the scale's neutral midpoint.[4] Table 1 shows the correlations and descriptive statistics for these items. With only one exception, the individual items included in the scale were strongly correlated. This exception, that the country "would be better off if we let in all immigrants who wanted to come," was also the only positively worded item included in the scale.

[4]We recoded the "Don't know" responses to the scale's midpoint for two reasons: (1) The "Don't know" response in this case is not substantially different than the midpoint on this item; (2) it assures we do not lose observations.

Measurement Invariance

As an initial test of reliability, we examined Cronbach's alpha across the 22 countries. Overall, the results indicate that the scale is reliable though there is also considerable variation across countries. In the United States, for example, α is 0.83, while in South Korea, α is only 0.53. A cursory glance at these reliability indicators suggests that the nativism scale may work better in some places than in others.

While Cronbach's alpha is a useful place to start, it assumes each individual item contributes the same amount to the underlying construct and that the error variance across items is equal and random (Ariely and Davidov, 2011). Tests of MIs that utilize multigroup factor analysis make neither of these assumptions. Specifically, we can represent the factor measurement model as follows:

$$\mathbf{y}_{ij} = \boldsymbol{\tau}_j + \boldsymbol{\Lambda}_j\, \eta_{ij} + \boldsymbol{\varepsilon}_{ij} \tag{1}$$

where subscript i means an individual respondent and j the specific group/country/time.[5] In the equation, $\mathbf{y}_{ij}$ is a column vector of the length of the number of question items measuring the latent variable. $\boldsymbol{\tau}_j$ and $\boldsymbol{\Lambda}_j$ denote the intercepts and factor loadings for group j; η_{ij} and $\boldsymbol{\varepsilon}_{ij}$ represent the latent (or called the common factor) score and item residuals of each individual, respectively (Kim et al., 2017). Scalar invariance, a strong sense of MI, means that the intercepts and factor loadings are identical across groups ($\boldsymbol{\Lambda}_j = \boldsymbol{\Lambda}_i, \boldsymbol{\tau}_j = \boldsymbol{\tau}_i,$ for $i \neq j$), so the researcher can be confident that any observed differences are caused by changes in latent variable scores rather than other systematic errors. When scalar invariance does not hold, groups or subjects over space and/or time respond differently to the items. As a result, factor means (or the latent variable means) cannot, or should not, be compared.[6]

Assessments of MI focus on four levels of invariance (Van de Schoot, Lugtig, and Hox, 2012). First, configural invariance indicates whether the latent variable is a function of the same set of observed variables. If so, this suggest that the structure behind the latent variable is the same across countries. Second, metric invariance indicates whether or not the factor loadings ($\boldsymbol{\Lambda}_j$) on the latent variable are the same across countries (Kim et al., 2017). Satisfying metric invariance indicates that the unit and the interval of the latent variable are equal across contexts, while failure to meet this level suggests that the construct does not have the same meaning across countries. In other words, a one unit increase in the latent variable would have a noticeably different effect across countries.

Third, scalar invariance is examined to evaluate the intercepts of each item variable ($\boldsymbol{\tau}_j$) across groups. Satisfying scalar invariance allows researchers to compare latent factor means, latent factor variances, and relevant covariance between groups (Kim et al., 2017). Failure to meet the scalar invariance indicates certain groups tend to systematically give higher or lower item responses. Fourth, strict invariance investigates if the residual variance of each item is consistent across groups after satisfying the assumptions of metric and scalar invariance. Strict invariance provides evidence that the mean differences across groups are driven by real group differences and not by error variance. While researchers prefer strict invariance as the ideal threshold for construct validity, they consider scalar invariance

[5]For ease of presentation, we are ignoring temporal differences for now. Also, the group in our context is one of the countries. We use the term "group" or "country" interchangeably in the following text. Our use of notation is consistent with Kim et al. (2017).

[6]For example, a study of religiosity reported that the frequency of attending religious services is not invariant (Davidov et al., 2014). In Muslim-dominant countries, women attend religious services less often but are not less religious.

empirically appropriate to compare factor or observed means because of the rigidness required to meet the strict invariance requirements (Davidov et al., 2008; Jang et al., 2017; Meredith, 1993).

Model Estimation

We used multigroup confirmatory factor analysis (MG-CFA) to test for configural, metric, and scalar invariance of the Ipsos nativism scale. MG-CFA is one of the most frequently used techniques for testing MI (Millsap, 2011; Van de Schoot et al., 2012). We did not test for strict invariance because prior research indicates that the standard is too rigorous to be met across a large number of countries (Byrne, 1994; Kim et al., 2017). MG-CFA starts with specifying a confirmatory factor analysis (CFA) model that reflects the theoretically operationalization of the construct. Separate CFA models are then fitted for each group, allowing comparisons across individual countries. We use an R package "lavaan" (latent variable analysis) to estimate our models and test for MI ⟨http://lavaan.ugent.be/⟩. "Lavaan" offers substantial flexibility to tune individual parameters within a complicated CFA or SEM framework (Rosseel, 2012) and is one of the most widely used R programming packages to conduct CFA or SEM studies.[7]

The model fit is interpreted based on three indices: chi-square, comparative fit index (CFI), and root mean squared error of approximation (RMSEA). Because chi-square values are sensitive to sample size and the number of groups included in the analysis (Bentler and Bonett, 1980; Jang et al., 2017), we primarily rely on CFI and RMSEA. By convention, the CFI ≥ 0.95 and the RMSEA ≤ 0.05 indicate acceptable model fits (Hu and Bentler, 1999). Within the MG-CFA context, however, previous research advocates for a more liberal RMSEA cutoff based on simulation results showing the number of groups affects the RMSEA cutoff (Rutkowski and Svetina, 2014). In this study, because we are investigating metric invariance across 19 countries, we use a 0.10 level threshold following the suggestions of prior studies (Jang et al., 2017; Rutkowski and Svetina, 2014).

Starting from the full scalar invariance model, we tuned the MI model using a top-down logic (Davidov et al., 2008; Meuleman et al., 2009; Reeskens and Hooghe, 2008). Recalling that full scalar invariance conveys that the intercepts and the factor loadings of items are the same across all countries ($\boldsymbol{\Lambda}_j = \boldsymbol{\Lambda}_i, \boldsymbol{\tau}_j = \boldsymbol{\tau}_i,\ \textit{for } i \neq j$), it is often considered a prerequisite for latent mean comparisons across countries or over time. However, it is not always necessary to achieve full scalar invariance, especially in the context of noisy, cross-country surveys. With partial scalar invariance, cross-national comparisons can be validated on the condition that at least two items per construct are equivalent (Byrne, Shavelson, and Muthén, 1989; Teo, 2015).

In terms of the fitting process, maximum likelihood estimation (MLE) is the conventional fitting scheme for SEM or CFA. However, MLE estimation assumes the multivariate normality of the input data (Curran, West, and Finch, 1996; Teo, 2015). Mardia's test of kurtosis and skewness indicates that the data here fail to meet this assumption (Korkmaz, Goksuluk, and Zararsiz, 2014; Mardia, 1970). In light of this finding, and to guarantee the robustness of our findings, we applied the maximum likelihood robust scaled (MLM) estimator to all of the CFA model estimations in this study (Rosseel, 2012).

[7]The R scripts are provided upon request.

Results

We began by computing a fully-constrained scalar invariance model that restricts all of the factor loadings and intercepts to be equivalent across all countries. Estimating this model resulted in a relatively poor model fit (CFI = 0.897; RMSEA = 0.138). Substantial improvements can be made, however, by freeing up parameters constrained by the assumptions of full metric invariance. We do this through an iterative process until there are no further modifications that would significantly improve the overall model fit. The results from this iterative process of model building are displayed in Table 2. As the modifications to the models proceed (indicated by each row in Table 2), the χ^2 decreases, reflecting constant improvements in model fit. The improvements eventually yield 21 models cumulatively with the last model achieving a good overall fit with CFI 0.953 and RMSEA 0.097. Because all countries have at least two items constrained to be identical, we can conclude that the index achieves partial scalar invariance. As a result, it is possible to make comparisons of the latent variable across countries. Substantively, this means the nativism scale is measuring the same latent construct across countries and that comparisons can be made across countries as to the level of nativism within each country.

The models presented in Table 2 may provide some insight into why the nativism scale might differ across countries. First, of the modifications necessary to achieve MI, only one is from an English-speaking country. In the current analysis, we lack sufficient evidence to know whether this reflects difference in translating question wordings or cultural differences in interpretation and response. Second, five countries (India, Brazil, Mexico, Argentina, and Spain) showed noninvariance on Q4, meaning that respondents in these countries did not interpret the question equivalently to respondents in other countries.[8] Notably, this item was the only item positively worded to gauge pro-immigration sentiments (i.e., support for allowing ALL immigrants to come into the country) and to minimize response bias. Third, the results from five countries (Sweden: 3.04, Germany: 2.92, Brazil: 2.17, Poland: 2.94, and Italy: 2.68, the rest of countries are 2.43) indicate a violation of the scalar invariance assumption on Q5. This item asked respondents whether their country would be stronger if immigration would be stopped. Interestingly, Sweden, which scores lowest on the overall nativism index, registers as more nativist on this one item. Responses to the question were from Sweden, which were polarized into strongly agree and strongly disagree categories. While a majority of Swedish respondents reject the idea that immigrants use too many government services, that immigrants take jobs from natives, or that natives should be given preference in hiring, significant percentages reported that the country would be better off if immigration were stopped. Our data do not speak directly to this point but preferences to halt immigration in this context may reflect concerns related to crime, personal security, or some other issue not gauged in this set of items (Pickering and Ham, 2017; Shechory-Bitton and Friedman, 2018).

Finally, we compared country rankings using the means for the nativism index (the average across the individual items) versus latent factor scores (from the MI model). We should note that the index averages and the latent factor scores are highly correlated ($r = 0.926$). As such, we are not making the case that latent means are preferable or that

[8]We make this determination by looking at both the factor loadings and the intercept for each individual question. For Q4, the parameters that needed to be unrestricted were for factor loadings (λ) and not the intercept (τ).

TABLE 2

Measurement Invariance Tests for Nativism—19 Groups, $N = 14{,}084$

Model	Description	χ^2	$\Delta\chi^2$	*df*	Δ *df*	CFI	CFI.Robust	RMSEA	RMSEA.Robust
Model 1	Full scalar invariance	3,618.051	NA	239	NA	0.897	0.897	0.138	0.138
Model 2	$+ \lambda_4^{India}$ free	3,285.08	332.971	238	1	0.907	0.907	0.131	0.131
Model 3	$+ \tau_5^{Sweden}$ free	3,126.036	159.044	237	1	0.912	0.912	0.128	0.128
Model 4	$+ \tau_5^{Germany}$ free	3,000.396	125.64	236	1	0.916	0.916	0.126	0.125
Model 5	$+ \lambda_4^{Brazil}$ free	2,903.258	97.138	235	1	0.919	0.919	0.124	0.124
Model 6	$+ \tau_5^{Brazil}$ free	2,793.836	109.422	234	1	0.922	0.922	0.121	0.121
Model 7	$+ \lambda_4^{Mexico}$ free	2,706.819	87.017	233	1	0.924	0.925	0.12	0.119
Model 8	$+ \lambda_1^{Germany}$ free	2,654.596	52.223	232	1	0.926	0.926	0.119	0.118
Model 9	$+ \tau_1^{Hungary}$ free	2,567.009	87.587	231	1	0.929	0.929	0.117	0.117
Model 10	$+ \tau_5^{Poland}$ free	2,484.053	82.956	230	1	0.931	0.932	0.115	0.115
Model 11	$+ \tau_3^{Poland}$ free	2,412.253	71.800	229	1	0.933	0.934	0.113	0.113
Model 12	$+ \tau_2^{Mexico}$ free	2,336.177	76.076	228	1	0.936	0.936	0.112	0.111
Model 13	$+ \tau_5^{Italy}$ free	2,275.038	61.139	227	1	0.937	0.938	0.11	0.110
Model 14	$+ \tau_3^{Belgium}$ free	2,208.146	66.892	226	1	0.939	0.94	0.109	0.108
Model 15	$+ \tau_1^{France}$ free	2,130.575	77.571	225	1	0.942	0.942	0.107	0.106
Model 16	$+ \tau_3^{France}$ free	2,050.51	80.065	224	1	0.944	0.945	0.105	0.104
Model 17	$+ \tau_1^{Canada}$ free	1,988.248	62.262	223	1	0.946	0.947	0.103	0.103
Model 18	$+ \tau_2^{Spain}$ free	1,928.163	60.085	222	1	0.948	0.948	0.102	0.101
Model 19	$+ \lambda_4^{Argentina}$ free	1,885.837	42.326	221	1	0.949	0.95	0.101	0.100
Model 20	$+ \lambda_4^{Spain}$ free	1,843.329	42.508	220	1	0.95	0.951	0.1	0.099
Model 21	$+ \tau_2^{Germany}$ free	1,770.129	73.2	219	1	0.953	0.953	0.098	0.097

NOTE: Each model (row) indicates what parameter (λ or τ) was freed from equivalence restrictions on top of prior models. For example, λ_4^{India} is the factor loadings of the question item 4 for India. In a similar way, the τ_5^{Sweden} refers to the intercepts of question item 5 for Sweden.

TABLE 3

Latent Mean (Mean Differences Compared to the Great Britain) Comparisons for Nativism—19 Countries, $N = 14{,}084$ (Ranked in Decreasing Order) and Additive Mean Comparisons

Latent_Mean_ Rank	Country	Latent_Mean	*SE*	*z*	*p*-Value	CI.Lower	CI.Upper	Mean_Additive	Additive_Mean_ Rank
1	Turkey	1.229	0.055	22.524	0.000	1.122	1.336	3.868	1
2	Hungary	1.004	0.069	14.598	0.000	0.870	1.139	3.548	3
3	India	0.845	0.05	16.825	0.000	0.746	0.943	3.344	9
4	Israel	0.845	0.058	14.653	0.000	0.732	0.957	3.653	2
5	France	0.704	0.053	13.338	0.000	0.601	0.808	3.374	7
6	Belgium	0.697	0.064	10.869	0.000	0.572	0.823	3.422	6
7	Argentina	0.666	0.064	10.474	0.000	0.541	0.79	3.373	8
8	South Africa	0.639	0.061	10.391	0.000	0.518	0.759	3.423	5
9	Italy	0.58	0.049	11.74	0.000	0.483	0.677	3.431	4
10	US	0.571	0.05	11.369	0.000	0.472	0.669	3.340	10
11	Australia	0.446	0.05	8.912	0.000	0.348	0.544	3.242	11
12	Spain	0.378	0.051	7.446	0.000	0.279	0.478	3.169	12
13	Mexico	0.159	0.057	2.783	0.005	0.047	0.271	2.937	18
14	Brazil	0.106	0.048	2.214	0.027	0.012	0.201	2.942	17
15	Canada	0.083	0.049	1.683	0.092	−0.014	0.18	2.976	16
16	Germany	0.064	0.055	1.162	0.245	−0.044	0.171	3.099	15
17	Poland	0.027	0.057	0.475	0.635	−0.085	0.14	3.131	14
18	Great Britain	0	0	NA	NA	0	0	3.165	13
19	Sweden	−0.14	0.064	−2.195	0.028	−0.264	−0.015	2.886	19

NOTE: The latent mean of Great Britain was set as zero for the purpose of model identification.

CFA is necessary when using the nativism index.[9] Still, there are interesting differences. As can be seen in Table 3, 13 of 19 countries have different rankings.

Conclusion

In this article, we utilized MG-CFA to establish the partial scalar invariance of the Ipsos nativism scale. Having achieved partial invariance, we have greater confidence that the scale yields useful insights in a comparative setting. Across 19 of the 22 countries included in this study, we can be confident that respondents are interpreting question items in comparable ways. Perhaps stated differently, the meaning of nativism does not significantly change from one country to the next. Having said this, future research should investigate those specific instances where the nativism index does not meet invariance assumptions. For example, the fact that attitudes were polarized in Sweden on questions of whether immigration should be halted is curious as Sweden scores very low on nativism overall. This might suggest the need for questions capturing nativist concerns relating to crime, personal security, or other issues not connected to services or jobs. Finally, this measurement exercise offers a number of conclusions regarding the level of nativism in mass publics; namely, the breadth of nativist sentiment differs depending on whether one utilizes an average of the observed variables or the latent factor scores.

REFERENCES

Aiken, L. S., J. A. Stein, and P. M. Bentler. 1994. "Structural Equation Analyses of Clinical Subpopulation Differences and Comparative Treatment Outcomes: Characterizing the Daily Lives of Drug Addicts." *Journal of Consulting and Clinical Psychology* 62(3):488.

Alemán, J., and D. Woods. 2016. "Value Orientations from the World Values Survey: How Comparable Are They Cross-Nationally?" *Comparative Political Studies* 49(8):1039–67.

Ariely, G., and E. Davidov. 2011. "Can We Rate Public Support for Democracy in a Comparable Way? Cross-National Equivalence of Democratic Attitudes in the World Value Survey." *Social Indicators Research* 104(2):271–86.

———. 2012. "Assessment of Measurement Equivalence with Cross-National and Longitudinal Surveys in Political Science." *European Political Science* 11(3):363–77.

Bentler, P. M., and D. G. Bonett. 1980. "Significance Tests and Goodness of Fit in the Analysis of Covariance Structures." *Psychological Bulletin* 88(3):588.

Byrne, B. M. 1994. "Testing for the Factorial Validity, Replication, and Invariance of a Measuring Instrument: A Paradigmatic Application Based on the Maslach Burnout Inventory." *Multivariate Behavioral Research* 29(3):289–311.

Byrne, B. M., R. J. Shavelson, and B. Muthén. 1989. "Testing for the Equivalence of Factor Covariance and Mean Structures: The Issue of Partial Measurement Invariance." *Psychological Bulletin* 105(3):456–66. Available at ⟨https://doi.org/10.1037/0033-2909.105.3.456⟩.

Carmines, E. G., M. J. Ensley, and M. W. Wagner. 2016. "Ideological Heterogeneity and the Rise of Donald Trump." Paper presented at the Forum, Inglewood, CA.

Curran, P. J., S. G. West, and J. F. Finch. 1996. "The Robustness of Test Statistics to Nonnormality and Specification Error in Confirmatory Factor Analysis." *Psychological Methods* 1(1):16–29. Available at ⟨https://doi.org/10.1037/1082-989X.1.1.16⟩.

[9] In theory, the latent mean has an advantage over a nativism index as it reflects a "purer" representation of the latent variable/construct, meaning that random error has been removed from the set of observed variables that comprise the scale and only the common variance remains (Aiken, Stein, and Bentler, 1994; Teo, 2015).

Davidov, E., and B. Meuleman 2012. "Explaining Attitudes Towards Immigration Policies in European Countries: The Role of Human Values." *Journal of Ethnic & Migration Studies* 38(5):757–75. Available at ⟨https://doi.org/10.1080/1369183X.2012.667985⟩.

Davidov, E., B. Meuleman, J. Billiet, and P. Schmidt. 2008. "Values and Support for Immigration: A Cross-Country Comparison." *European Sociological Review* 24(5):583–99.

Davidov, E., B. Meuleman, J. Cieciuch, P. Schmidt, and J. Billiet. 2014. "Measurement Equivalence in Cross-National Research." *Annual Review of Sociology* 40:55–75.

González, F. E., and C. Young. 2017. "The Resurgence and Spread of Populism?" *SAIS Review of International Affairs* 37(1):3–18.

Gratton, B. 2018. "Race or Politics? Henry Cabot Lodge and the Origins of the Immigration Restriction Movement in the United States." *Journal of Policy History* 30(1):128–57. Available at ⟨https://doi.org/10.1017/S0898030617000410⟩.

Hatton, T. J. 2016. "Immigration, Public Opinion and the Recession in Europe." *Economic Policy* 31(86): 205–46.

Hellström, A., and P. Hervik. 2014. "Feeding the Beast: Nourishing Nativist Appeals in Sweden and in Denmark." *Journal of International Migration and Integration* 15(3):449–67.

Hooghe, M., and R. Dassonneville. 2018. "Explaining the Trump Vote: The Effect of Racist Resentment and Anti-Immigrant Sentiments." *Political Science & Politics* 1–7.

Hu, L. T., and P. M. Bentler. 1999. "Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria Versus New Alternatives." *Structural Equation Modeling: A Multidisciplinary Journal* 6(1):1–55.

Inglehart, R., and P. Norris. 2017. "Trump and the Populist Authoritarian Parties: The Silent Revolution in Reverse." *Perspectives on Politics* 15(2):443–54.

Ivarsflaten, E. 2005. "Threatened by Diversity: Why Restrictive Asylum and Immigration Policies Appeal to Western Europeans." *Journal of Elections, Public Opinion and Parties* 15(1):21–45. Available at ⟨https://doi.org/10.1080/13689880500064577⟩.

Jang, S., E. S. Kim, C. Cao, T. D. Allen, C. L. Cooper, L. M. Lapierre, M. P. O'Driscoll, J. I. Sanchez, P. E. Spector, S. A. Y. Poelmans, N. Abarca, M. Alexandrova, A.-S. Antoniou, B. Beham, P. Brough, I. Carikci, P. Ferreiro, G. Fraile, S. Geurts, U. Kinnunen, C.-q. Lu, L. Lu, I. F. Moreno-Velázquez, M. Pagon, H. Pitariu, V. Salamatov, O.-l. Siu, S. Shima, M. K. Schulmeyer, K. Tillemann, M. Widerszal-Bazyl, and J.-M. Woo. 2017. "Measurement Invariance of the Satisfaction with Life Scale Across 26 Countries." *Journal of Cross-Cultural Psychology* 48(4):560–76.

Johnson, T. P. 1998. "Approaches to Equivalence in Cross-Cultural and Cross-National Survey Research." *ZUMA Nachrichten Spezial* 3:1–40.

Kim, E. S., C. Cao, Y. Wang, and D. T. Nguyen. 2017. "Measurement Invariance Testing with Many Groups: A Comparison of Five Approaches." *Structural Equation Modeling: A Multidisciplinary Journal* 24(4):524–44.

King, G., C. J. Murray, J. A. Salomon, and A. Tandon. 2004. "Enhancing the Validity and Cross-Cultural Comparability of Measurement in Survey Research." *American Political Science Review* 98(1):191–207.

Knoll, B. R. 2013. "Implicit Nativist Attitudes, Social Desirability, and Immigration Policy Preferences Implicit Nativist Attitudes, Social Desirability, and Immigration Policy Preferences." *International Migration Review* 47(1):132–65. Available at ⟨https://doi.org/10.1111/imre.12016⟩.

Korkmaz, S., D. Goksuluk, and G. Zararsiz. 2014. "MVN: An R Package for Assessing Multivariate Normality." *R Journal* 6(2):151–62.

LeVine, R. A., and D. T. Campbell. 1972. *Ethnocentrism: Theories of Conflict, Ethnic Attitudes, and Group Behavior*. Oxford, England: John Wiley & Sons.

MacWilliams, M. C. 2016. "Who Decides When the Party Doesn't? Authoritarian Voters and the Rise of Donald Trump." *Political Science & Politics* 49(4):716–21.

Mardia, K. V. 1970. "Measures of Multivariate Skewness and Kurtosis with Applications." *Biometrika* 57(3):519–30. Available at ⟨https://doi.org/10.1093/biomet/57.3.519⟩.

Meredith, W. 1993. “Measurement Invariance, Factor Analysis and Factorial Invariance.” *Psychometrika* 58(4):525–43.

Meuleman, B., E. Davidov, and J. Billiet. 2009. “Changing Attitudes Toward Immigration in Europe, 2002–2007: A Dynamic Group Conflict Theory Approach.” *Social Science Research* 38(2):352–65.

Millsap, R. E. 2011. *Statistical Approaches to Measurement Invariance*. New York, NY: Routledge/Taylor & Francis Group.

Mudde, C. 2007. *Populist Radical Right Parties in Europe*, Vol. 22. Cambridge: Cambridge University Press.

Mutz, D. C. 2018. “Status Threat, Not Economic Hardship, Explains the 2016 Presidential Vote.” Proceedings of the National Academy of Sciences. Available at ⟨https://www.pnas.org/content/115/19/E4330⟩.

Newman, B. J., S. Shah, and L. Collingwood. 2018. “Race, Place, and Building a Base: Latino Population Growth and the Nascent Trump Campaign for President.” *Public Opinion Quarterly* 82(1):122–34.

Oberski, D. L. 2014. “Evaluating Sensitivity of Parameters of Interest to Measurement Invariance in Latent Variable Models.” *Political Analysis* 22(1):45–60.

Oberski, D. L., J. K. Vermunt, and G. B. Moors. 2015. “Evaluating Measurement Invariance in Categorical Data Latent Variable Models with the EPC-Interest.” *Political Analysis* 23(4):550–63.

Pickering, S., and J. Ham. 2017. *The Routledge Handbook on Crime and International Migration*. Abingdon, UK: Routledge.

Reeskens, T., and M. Hooghe. 2008. “Cross-Cultural Measurement Equivalence of Generalized Trust. Evidence from the European Social Survey (2002 and 2004).” *Social Indicators Research* 85(3):515–32.

Rosseel, Y. 2012. “Lavaan: An R Package for Structural Equation Modeling and More. Version 0.5–12 (BETA).” *Journal of Statistical Software* 48(2):1–36.

Rutkowski, L., and D. Svetina. 2014. “Assessing the Hypothesis of Measurement Invariance in the Context of Large-Scale International Surveys.” *Educational and Psychological Measurement* 74(1):31–57.

Schaffner, B. F., M. MacWilliams, and T. Nteta. 2017. “Hostile Sexism, Racism Denial, and the Historic Education Gap in Support for Trump.” P. 99 in *2016 Presidential Election: The Causes and Consequences of a Political Earthquake*.

Shechory-Bitton, M., and E. S. Friedman. 2018. “Fear of Crime and Terrorism Among Israeli and Swedish Citizens.” Pp. 577–95 in *Refugees and Migrants in Law and Policy*. Berlin, Germany: Springer.

Sides, J., M. Tesler, and L. Vavreck. 2017. “The 2016 US Election: How Trump Lost and Won.” *Journal of Democracy* 28(2):34–44.

Teo, T. 2015. “Comparing Pre-Service and In-Service Teachers’ Acceptance of Technology: Assessment of Measurement Invariance and Latent Mean Differences.” *Computers & Education* 83:22–31.

Van de Schoot, R., P. Lugtig, and J. Hox. 2012. “A Checklist for Testing Measurement Invariance.” *European Journal of Developmental Psychology* 9(4):486–92.

Welzel, C., and R. F. Inglehart. 2016. “Misconceptions of Measurement Equivalence: Time for a Paradigm Shift.” *Comparative Political Studies* 49(8):1068–94.

Young, C. 2016. “It’s Nativism: Explaining the Drivers of Trump’s Popular Support.” Available at ⟨https://www.ipsos.com/sites/default/files/2016-09/Its_Nativism_Explaining_the_Drivers_of_Trumps_Popular_Support.pdf⟩.