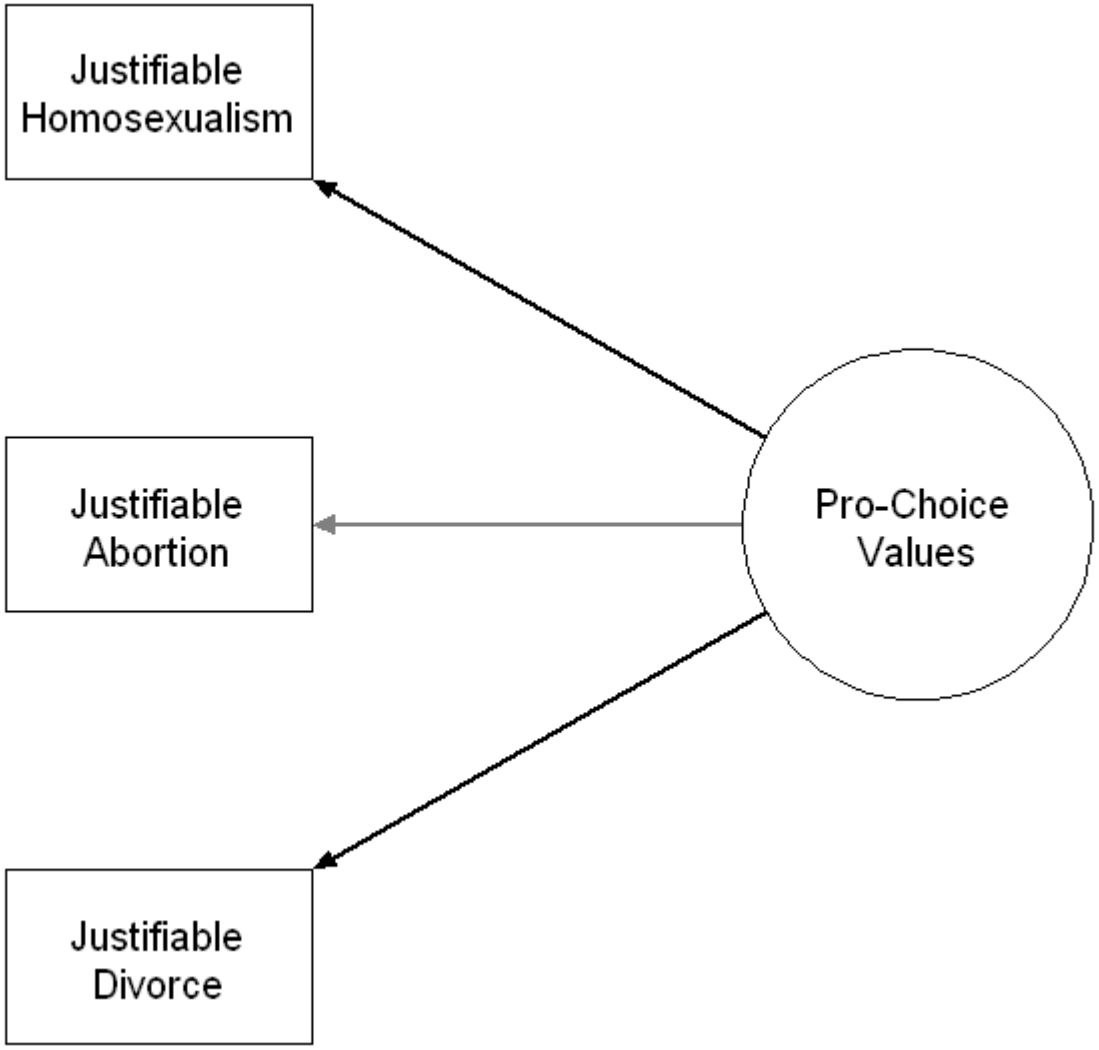
Establishing invariance of pro-choice values across countries in the World Values Survey: an approximate Bayesian approach

Boris Sokolov

Laboratory for Comparative Social Research
Higher School of Economics

April 18, 2015
St. Petersburg

Pro-Choice Values (all items are ten-point scales)



Motivation

- Recent debate on the issue of measurement invariance, and broadly, measurement validity of emancipative values (Aléman and Woods 2015; Welzel and Inglehart 2016): comparability of values across the world is under question.
- In contrast to other components of the emancipative values, pro-choice values seem to be more or less robust cross-nationally
- Pro-choice values may in some contexts be as powerful explanatory variable as the emancipative values are (Inglehart , Puranen, and Welzel 2015; Sokolov LCSR conference 2015)

CFA of 12 variables for ten cultural zones (WVS, 6th wave)

Variable	Factor	Islamic East	Indic East	Sinic East	Orthodox East	Old West	Reformed West	New West	Returned East	Latin America	sub-Saharan Africa
Independence	Autonomy	**0.256**	0.639	**0.137**	0.404	0.558	0.444	**0.280**	0.787	**0.316**	0.533
Imagination	Autonomy	0.325	**0.193**	0.439	**0.249**	0.436	0.494	0.401	0.413	**0.297**	0.342
Obedience	Autonomy	0.915	0.327	0.457	0.812	0.588	0.619	0.465	**0.079**	0.406	0.625
Jobs	Equality	0.545	0.347	0.339	0.368	**0.246**	**–**	**-0.116**	**–**	**0.118**	0.338
Leaders	Equality	0.770	0.773	0.727	0.705	0.886	0.749	0.745	0.642	0.877	0.881
Education	Equality	**0.292**	0.528	0.720	0.569	0.640	0.864	0.794	0.704	0.590	0.441
Homosexuality	**Choice**	**0.623**	**0.707**	**0.793**	**0.515**	**0.831**	**0.806**	**0.875**	**0.651**	**0.774**	**0.813**
Abortion	**Choice**	**0.874**	**0.760**	**0.664**	**0.766**	**0.714**	**0.783**	**0.722**	**0.793**	**0.557**	**0.903**
Divorce	**Choice**	**0.456**	**0.745**	**0.786**	**0.739**	**0.665**	**0.832**	**0.685**	**0.786**	**0.704**	**0.670**
Speech	Voice	**0.245**	**–**	**0.283**	**–**	0.325	**–**	**0.121**	**–**	–	-0.337
Say_nat	Voice	**0.268**	**n.s.**	0.449	0.409	**0.180**	0.447	**-0.109**	0.645	0.582	0.849
Say_local	Voice	0.498	**n.s.**	0.559	0.431	0.587	0.711	0.302	**0.260**	0.406	**0.214**
Autonomy	EVI	0.416	**0.145**	0.597	**n.s.**	**0.296**	0.704	0.832	0.554	0.460	0.562
Equality	EVI	0.505	**0.127**	0.525	0.304	0.384	0.614	0.560	0.419	**0.229**	**0.066**
Choice	EVI	0.340	1.000	0.630	0.686	0.768	0.616	0.803	0.781	0.769	0.651
Voice	EVI	0.785	**n.s.**	0.572	0.449	0.785	**0.260**	0.977	**n.s.**	0.336	**0.246**

Notes: Entries are standardized factor loadings. All estimates are significant at 0.05 level (except those marked as n.s. = non-significant). Loadings in bold are those lower than 0.32. Variable intercepts, thresholds and variances are not shown.

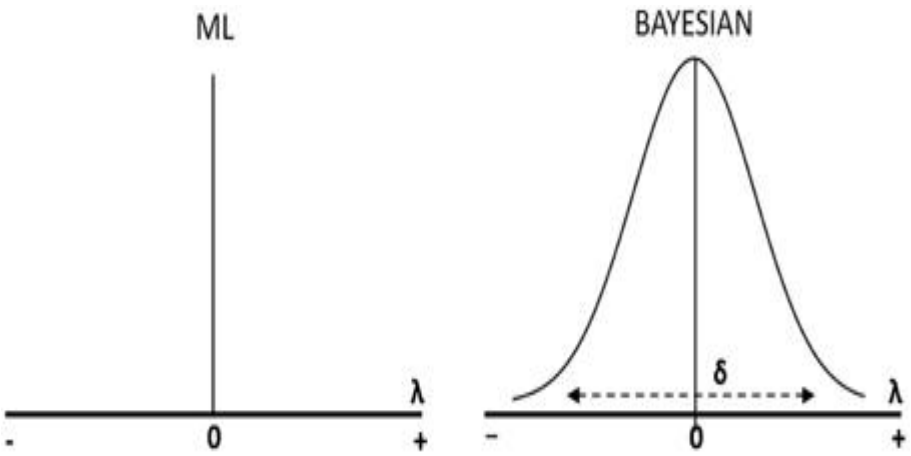
Measurement Invariance

- Measurement invariance "is a property of a measurement instrument (in the case of survey research: a questionnaire), implying that the instrument measures the same concept in the same way across various subgroups of respondents" (Davidov et al. 2014, p. 58)
- Establishing measurement invariance is the necessary requirement for cross-national comparisons of latent means.
- Formally, the assumption of measurement invariance implies that key model parameters (for CFA these are factor loadings and indicator intercepts) should be the same in each relevant subsample.
- **Why Bayesian approach?** Classical exact invariance testing is too strict. Even cross-zone invariance of pro-choice values is rejected, using exact approach. Bayesian approximate approach is more tolerant to small country-specific deviations from the sample averaged parameters.
- In the Bayesian approach the differences of factor loadings or intercepts between groups are assumed to be approximately zero with a mean of zero and some small variance (Van de Shoot et al. 2013; Zercher et al. 2015)

The Concept of Approximate Invariance

- The trick is to permit “small” differences between group-specific parameters instead of fixing those differences exactly to zero
- The model fit can be evaluated based on the posterior predictive p-value (PPP) and the confidence interval (CI) for the difference between the observed and replicated chi-square values (those including zero indicate well-fitting models). PPP should be higher than 0.05 and, in ideal, close to 0.5.

Difference in parameter estimation between Maximum Likelihood (ML) and the Bayesian approach (see van de Schoot et al., 2013; Zercher et al. 2015).



Measurement Models

Approximate Invariance Model

- For the item "**Justifiable: Divorce**" the factor loading is fixed to 1 and the intercept is fixed to 0 in all groups for achieving identification.
- For the items "Justifiable: Abortion" and "Justifiable: Homosexualism" group-specific factor loadings and intercepts are allowed to deviate from the global mean with group-specific deviations distributed as $\sim N(0, 0.05)$

Partial Approximate Invariance Models

- For the item "Justifiable: Divorce" the factor loading is fixed to 1 and the intercept is fixed to 0 in all groups for achieving identification.
- For the item "Justifiable: Abortion" group-specific factor loadings and intercepts are allowed to deviate from the global mean with deviations distributed as $\sim N(0, 0.05)$
- For the item "Justifiable: Homosexualism" both group-specific loadings and intercepts are freely estimated.

Samples

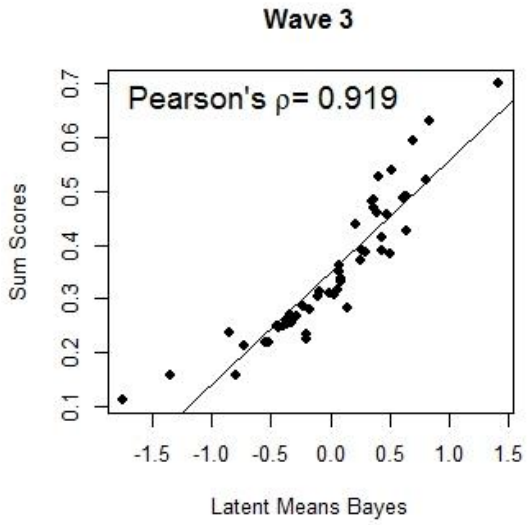
- Wave 1 (1981-1984): 8 countries, 10307 respondents
- Wave 2 (1989-1993): 18 countries, 24558 respondents
- Wave 3 (1994-1998): 51 countries, 72964 respondents
  - Bangladesh, Pakistan, and Turkey excluded (because one or more relevant questions were not asked in those countries)
- Wave 4 (1999-2004): 37 countries, 55454 respondents
  - Turkey, Morocco and Iraq excluded
- Wave 5 (2005-2009): 54 countries, 75523 respondents
  - Iraq, Morocco, Peru and Egypt excluded
- Wave 6 (2010-2014): 58 countries, 83446 respondents
  - Kuwait and Egypt excluded

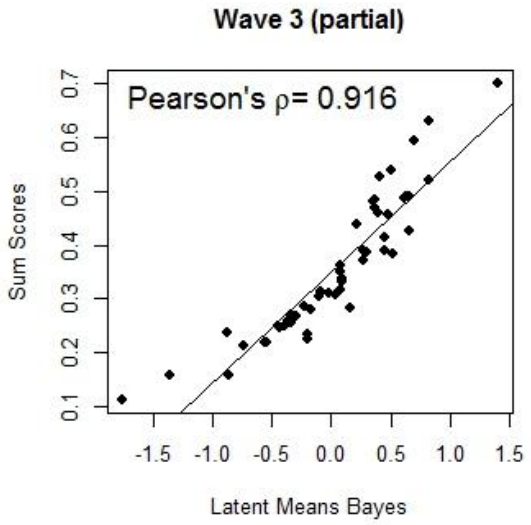
- Full information maximum likelihood (FIML) is used to deal with missing values

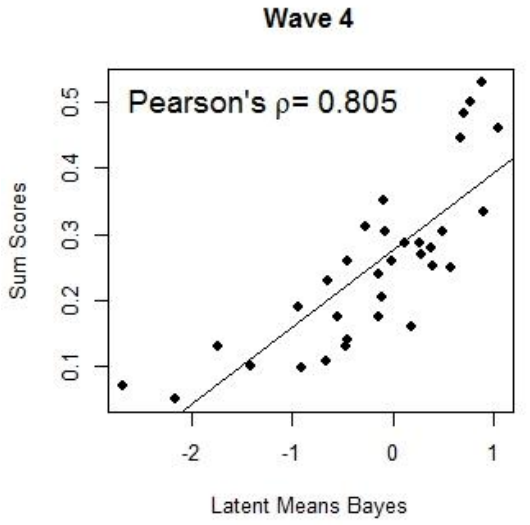
Results

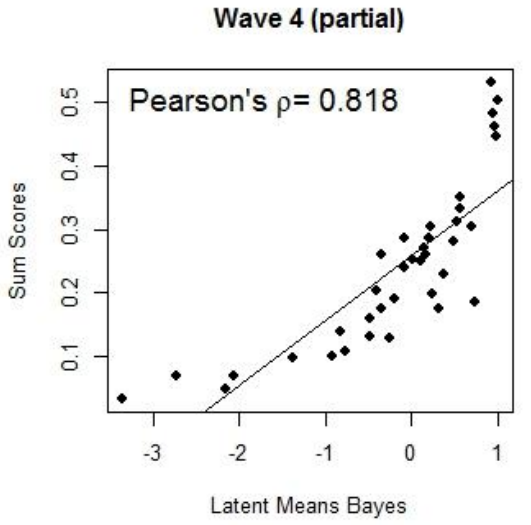
Wave	Approximate Invariance Model		Partial Approximate Invariance Model		Notes
	PPP	χ^2 Confidence Interval	PPP	χ^2 Confidence Interval	
Wave 1	0.453	-31.787; 35.615	–	–	
Wave 2	0.349	-40.172; 61.866	–	–	
Wave 3	0.266	-57.416; 114.544	0.344	-65.862; 104.924	
Wave 4	**0.016**	**6.559; 155.751**	**0.045**	-10.403; 140.402	
Wave 4 reduced	0.210	-42,872; 99.932	0.345	-57.998; 83.172	33 countries: Saudi, Bangladesh, Pakistan and Algeria excluded
Wave 5	0.106	-32.144; 140.523	0.159	-44.015; 133.167	
Wave 6	**0.034**	-5.370; 179.934	0.081	-24.731; 158.249	
Wave 6 reduced	0.139	-32.289; 127.222	0.217	-46.739; 113.302	46 Countries: Islamic East excluded

Pro-choice values: Relationship between sum scores and scores based on the Bayesian estimation in Waves 3 and 4.

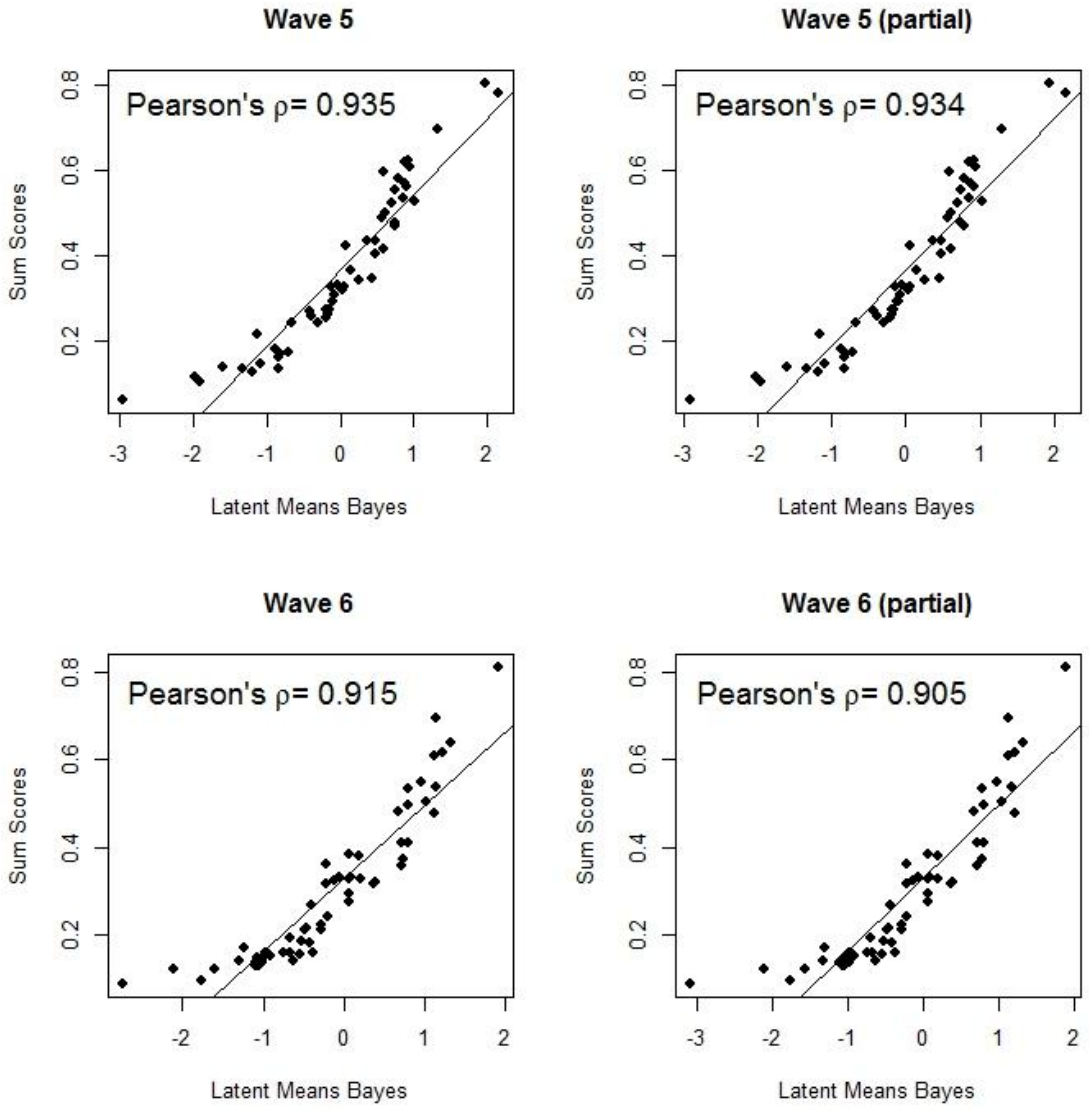








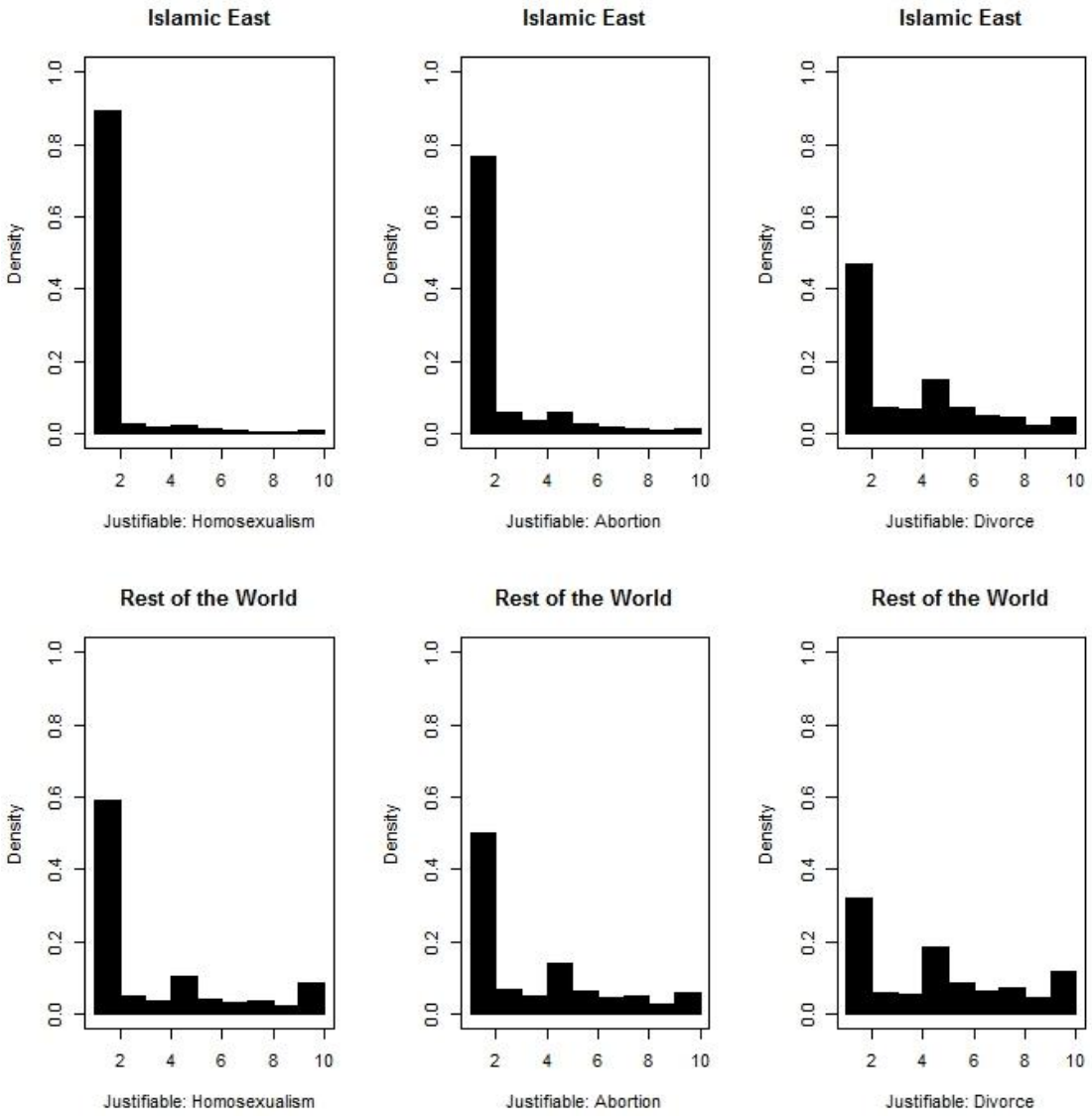
Pro-choice values: Relationship between sum scores and scores based on the Bayesian estimation in Waves 5 and 6.



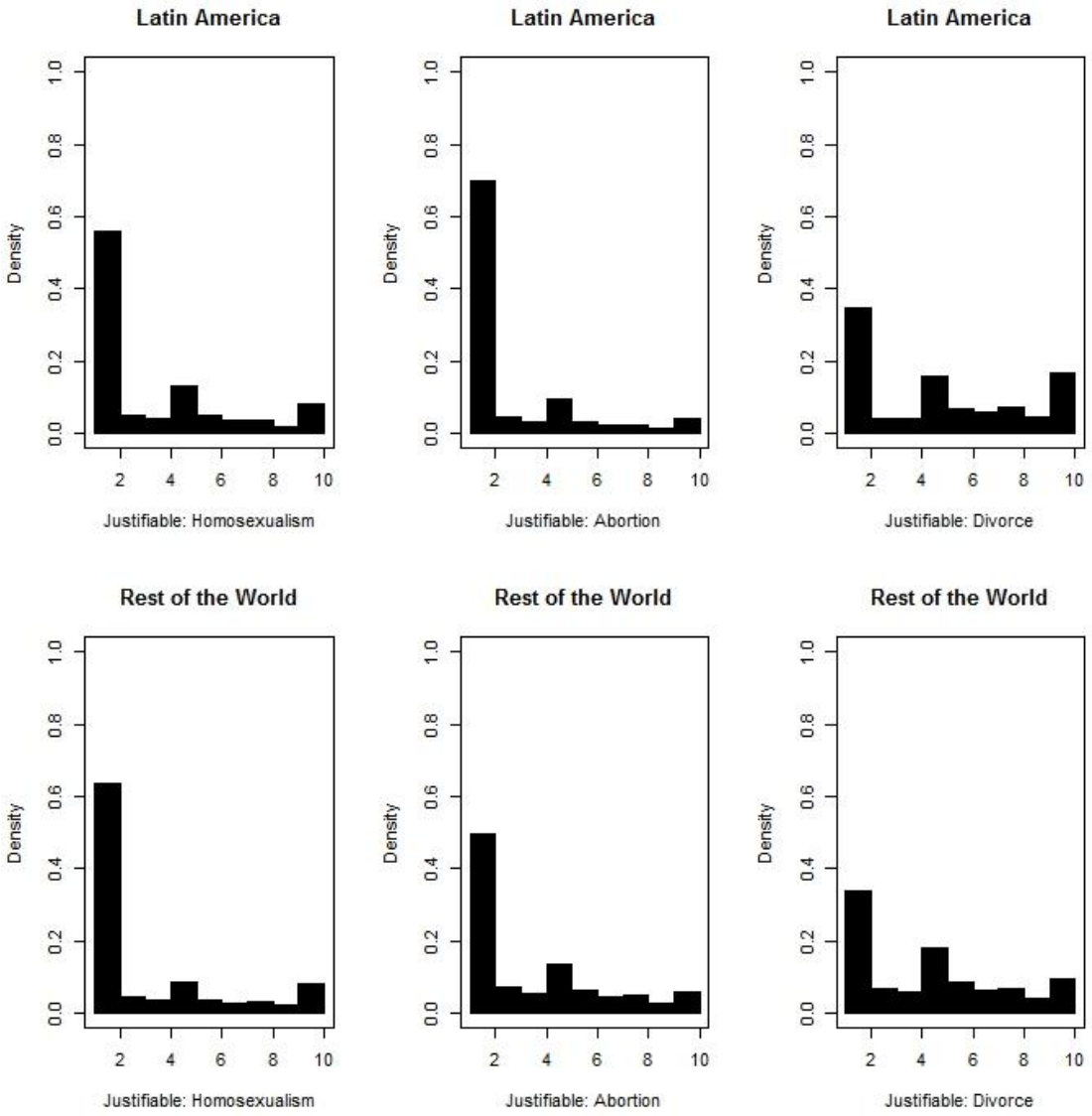
The most deviating countries in two problematic waves (Group-Specific PPPs lower than 0.3)

- Wave 4: Algeria, Bangladesh, Pakistan, Saudi Arabia, Argentina, Peru, Puerto Rico, Venezuela
- Wave 6: Palestine, Ghana, Jordan, Lebanon, Morocco, Pakistan, and Tunisia

Individual responses on the relevant survey items: Islamic East vs. Rest of the World



Individual responses on the relevant survey items: Latin America vs. Rest of the World



Summary of main findings

- Partial approximately invariance model fits well in each WVS wave
- Full Approximately Invariant Model fits well only in Waves 1, 2, 3, and 5.
- Model fit depends on the choice of the marker variable
- The most variable item is "***Whether Homosexualism is Justifiable***"?
- The most deviating countries are Muslim countries.
- Large country-specific response biases in the Muslim world can be nevertheless qualified as an example of the "elevator effect" (Welzel 2013; Welzel and Inglehart 2016):

  low society-level tolerance to deviant forms of reproductive behaviour in Muslim countries causes downward bias in individual responses (kind of social desirability bias)

❖ Aggregated scores on pro-choice values can (a) be validly compared across WVS countries and also (b) used in causal analyses relating these scores to other social and political country-level outcomes.

Further steps

- Testing within-country temporal invariance of pro-choice values
- Testing stricter forms of invariance (i.e. assuming smaller prior variances for between-countries differences in parameters' magnitudes)
- Testing sensitivity of previous findings exploiting the concept of pro-choice values: are they robust to measurement model correction?
- Theoretical work explaining the consistence of pro-choice values across the world

Thank you for attention!

This report was presented at the 6th LCSR International Workshop
“Trust, Social Capital and Values in a Comparative Perspective”,
which held within the XVII April International Academic Conference on Economic and Social Development.

April 18 – April 22, 2016 - Higher School of Economics, Moscow.

https://lcsr.hse.ru/en/seminar2016

Настоящий доклад был представлен на VI международном рабочем семинаре ЛССИ
«Доверие, социальный капитал и ценности в сравнительной перспективе»,
прошедшего в рамках XVII Апрельской международной научной конференции НИУ ВШЭ «Модернизация экономики и общества».

18 – 22 апреля, 2016 – НИУ ВШЭ, Москва.

https://lcsr.hse.ru/seminar2016