

## Technical Appendix C: Methods

### Multilevel Regression Models

As not all readers may be familiar with the analytical methods used in this study, a brief note helps to clarify the techniques. The firewall theory developed in chapter 1 and 2 predicts that individual use of the news media has a *direct* effect on individual values. In addition, the account predicts that a *cross-level interaction* effect will also be apparent, as external and internal barriers to information flows in each society will interact with individual patterns of news media use.

To operationalize these factors, the key models in the second section of the book involve measurement at two distinct levels. A representative sample of individual respondents (level 1) is nested within national-level contexts (level 2). The World Values Survey was conducted among a representative random sample of the adult population within each nation-state.

The danger of using Ordinary Least Squares (OLS) regression for analysis is that the standard errors of the regression coefficients can be inaccurate for contextual and cross-level variables, by overestimating the degrees of freedom, and therefore tests of significance can prove misleading. OLS models can seek to control for national variations by using a pooled model including dummy variables for each country, but this becomes inefficient with the coverage of many nations. Alternatively OLS models can be run with no pooling, where separate models are run for each nation or type of media environment, but this is also clumsy.

Given the use of multilevel data, hierarchical linear models (HLM) are most appropriate for analysis, including multilevel regression analysis.<sup>1</sup> The models in this study use restricted maximum likelihood techniques (REML) to estimate direct and cross-level effects for hierarchical data. Individual respondents are thus grouped into nation-states. Each nation-state has a different set of parameters for the random factors, allowing intercepts and slopes to vary by nation.<sup>2</sup>

Level 1 in our core models includes the following *individual-level* measures: male gender (0/1), household income using a 10-point scale, age (in years), the education scale, and the media use 5-point scale (or each of the separate dummy variables for use of newspapers, radio/TV and the Internet).

Level 2 includes the following *national-level* variables: the standardized KOF globalization index, the standardized Freedom House press freedom index, and the standardized level of economic development (per capita GDP (2006) in Purchasing Power Parity). As each of these were highly inter-

correlated ( $R=.640$  and above), the standardized scores were combined into the Cosmopolitanism Index. Each component was equally weighted. In calculating the Cosmopolitanism Index, the appropriate year of these indices was matched to the closest year of the national survey.

*Cross-level* interactions are also included, as the effects of media use are expected to be moderated by the level of cosmopolitanism within each country. To measure the joint effects of media use at individual level, while taking account of the environment at national-level, models include the Cosmopolitanism Index\*media use.

All variables are described in Technical Appendix A and Table C.1. In SPSS 16.0 Mixed Models, the iterative restricted maximum likelihood (REML) algorithm was used for estimating parameters. In hierarchical linear models, as is customary, all independent variables were centered, by subtracting the grand mean (which becomes zero). The standardized independent variables all have a standard deviation of 1.0. This process also helps to guard against problems of collinearity in the independent variables in the OLS models. The dependent variables are all converted into 100-point scales for ease of comparison across different tables. The independent variables were treated as fixed components, reflecting the weighted average for the slope across all groups, while nation was treated as a random component, capturing the country variability in the slope. The strength of the beta coefficients (slopes) can be interpreted intuitively as how much change in the dependent variable is generated by a one-percent change in each independent variable.

The treatment of missing data is also important. Mean substitution replaced missing data for individual-level income where this was omitted in the national surveys conducted in two countries (Argentina and Jordan). The country coverage of the KOF Globalization Index is skewed towards more developed societies and some developing countries are omitted, due to problems of missing data from official statistics. Accordingly the regional mean was used to estimate the position for seven countries which were not included in the 2008 KOF Globalization Index (Andorra, Taiwan, Ethiopia, Moldova, Viet Nam, Burkina Faso, and Serbia). Models were tested with and without these treatments to check that they did not have a substantial effect on the interpretation of the results.

The multilevel regression models used in this study usually generate small differences in the size of the slope coefficient ( $b$ ) compared with the results of OLS models, but the average standard errors for level 2 variables often tend to be slightly larger. The process is thus more rigorous and conservative, avoiding Type I errors (false positives, concluding that a statistically-significant difference exists when, in truth,

there is no statistical difference). The goodness of fit statistic in OLS is the adjusted  $R^2$ , where models with a higher coefficient indicate that it accounts for more of the variance. In the REML model, by contrast, Schwarz's Bayesian Criterion (BIC) is used, where the model with the lower value is the best fitting.

Table C.2 below compares the results of using both the OLS and the REML models, where the 100-point 'trust in outsiders' scale is used illustratively as the dependent variable. When the three components of the Cosmopolitanism index were entered separately into the OLS model, the results encountered severe problems of multicollinearity (using the indicators of tolerance and Variance Inflation Factor), making the coefficients unreliable and the signs inconsistent. The use of the combined Cosmopolitanism Index avoids potential problems of multicollinearity. Comparison of the estimates generated by the OLS and REML models in Table C.2 show that most of the estimates of the slope and intercept are very similar. The main contrast is that the OLS model can inflate the appropriate degrees of freedom at national level, whereas the REML model is preferable by providing a more rigorous and conservative estimate of significance at national level.

Both models confirm that, even with a battery of social controls, *media use at an individual level has a direct effect on trust in outsiders*. Each percentage point increase in the media use scale corresponds, on average, to a 1.5% increase in trust in outsiders, when all independent variables are set equal to their grand mean values. Trust in outsiders is also significantly increased by age, income and education (but not gender). The comparison of individual-level factors shows that media use had a stronger effect on trust in outsiders than any of the other individual-level factors.

At national level, the REML model suggests that the Cosmopolitanism Index by itself has a very strong and significant direct effect on trust in outsiders; each percentage point increase in the Cosmopolitanism Index corresponds, on average, to a 2.0% increase in trust in outsiders, when all independent variables are set equal to their grand mean values. The direct experience of living in a cosmopolitan society therefore generates more trust in outsiders than use of the media alone.

Moreover the models also suggest that, over and above these effects, a cross-level interaction effect is also present: *the positive effect of media use on trust in outsiders is further increased in the most cosmopolitan societies*. As discussed further in chapter 6, the results of the analysis therefore confirm the firewall theory.

The estimates of covariance suggest that the national intercepts were significant and strong ( $b=.65.9$ ,  $s.e.14.1$ ,  $p.000$ ), capturing the variability in the trust in outsiders scale among countries.

Excluding the insignificant predictors generated the following equations. In the notation,  $_1$  refers to level 1 (individual) and  $_2$  to level 2 (national) variables.

#### **Model A: OLS Regression analysis without the interaction terms**

$$Y_{\text{TRUST}} = 55.339 + 975x_{\text{AGE1}} + .766x_{\text{INCOME1}} + 1.345x_{\text{EDUCATION1}} + 1.488x_{\text{MEDIAUSE1}} + 2.405x_{\text{COSMOPOLITANISM2}} + .725x_{\text{COSMOPOLITANISM2}*\text{MEDIAUSE1}}$$

#### **Model B: REML Multilevel Regression analysis**

$$Y_{\text{TRUST}} = 55.096 + 1.45x_{\text{AGE1}} + .628x_{\text{INCOME1}} + 2.177x_{\text{EDUCATION1}} + 1.583x_{\text{MEDIAUSE1}} + 2.065x_{\text{COSMOPOLITANISM2}} + .426x_{\text{COSMOPOLITANISM2}*\text{MEDIAUSE1}}$$

**Table C.1: Description of the core independent variables, WVS 2005-7**

	N	Unstandardized			Standardized (z scores)		
		Min	Max	Mean	Min	Max	Mean
<b>INDIVIDUAL-LEVEL</b>							
<b>Demographic characteristics</b>							
Age (years)	78,017	16	89	42	-1.55	2.86	.000
Male Gender	78,320	0	1	.48	-.96	1.04	.000
<b>Socio-economic resources</b>							
Household Income scale	78,416	1	10	4.7	-1.67	2.42	.000
Education scale	77,822	1	9	5.3	-1.84	1.55	.000
<b>Media use</b>							
Media use 100-point scale	65,729	0	100	38.5	-1.16	1.85	.000
Read newspaper (0/1)	70,692	0	1	.57	-1.15	.86	.000
Use radio/TV news(0/1)	70,306	0	1	.88	-2.72	.37	.000
Use internet/email(0/1)	69,037	0	1	.29	-.65	1.53	.000
<b>NATIONAL-LEVEL</b>							
External barriers: Globalization index		0	100	62	-2.17	1.81	.000
Internal barriers: Media Freedom		0	100	58	-1.83	1.40	.000
Development: GDP per capita in PPP		\$155	\$40,947	\$10,499	-.82	2.40	.000
Cosmopolitanism index (ZGlobalization+ZFreedom+ZGDP)		-4.81	5.03	.010	-1.81	1.88	.000
<b>CROSS-LEVEL INTERACTIONS</b>							
Globalization*media use scale		0	9002	2606	-1.07	2.62	.000
Media freedom*media use scale		0	9100	2558	-1.01	2.58	.000
Cosmopolitanism index*media use		-8.91	9.32	.95	-3.88	3.29	.000
Valid N Listwise	61,513						

Source: World Values Survey 2005-7.

**Table C.2: Comparison of OLS and multilevel regression models explaining trust in outsiders**

	Model A: OLS regression				Model B: REML Multilevel regression		
	b	SE	Beta	Sig.	b	SE	Sig.
<b>INDIVIDUAL-LEVEL</b>							
<b>Demographic characteristics</b>							
Age (years)	<b>.975</b>	<b>.101</b>	<b>.041</b>	<b>.000</b>	<b>1.45</b>	<b>.099</b>	<b>.000</b>
Gender (male=1)	-.022	.094	-.000	.815	-.099	.090	.275
<b>Socio-economic resources</b>							
Household Income 10-pt scale	<b>.766</b>	<b>.101</b>	<b>.032</b>	<b>.000</b>	<b>.628</b>	<b>.101</b>	<b>.000</b>
Education 9-pt scale	<b>1.345</b>	<b>.114</b>	<b>.055</b>	<b>.000</b>	<b>2.177</b>	<b>.119</b>	<b>.000</b>
<b>Media use</b>							
Media use 100-point scale	<b>1.488</b>	<b>.115</b>	<b>.062</b>	<b>.000</b>	<b>1.583</b>	<b>.112</b>	<b>.000</b>
<b>NATIONAL-LEVEL</b>							
Cosmopolitanism index	<b>2.405</b>	<b>.042</b>	<b>.257</b>	<b>.000</b>	<b>2.065</b>	<b>.459</b>	<b>.000</b>
<b>CROSS-LEVEL INTERACTIONS</b>							
Cosmopolitanism*media use scale	<b>.725</b>	<b>.038</b>	<b>.077</b>	<b>.000</b>	<b>.312</b>	<b>.039</b>	<b>.000</b>
Constant	55.339				55.096		
Goodness of fit: Adjusted R <sup>2</sup>	.121						
Goodness of fit: Schwarz's BIC					493,479		
N. respondents	56,413				55,090		
N. nations	44				44		

**Note:** All independent variables have been standardized between 0-1 using mean centering (z-scores). *Model A* presents the results of OLS regression models while *Model B* presents the results of the REML multilevel regression models. The trust in outsiders (other nationalities and other religions) 100-point scale was the dependent variable. The 100-point media use scale combined use of newspapers, radio/TV, the internet, books, and magazines for information. *Model A* reports the unstandardized beta coefficients (b), the standard errors, the standardized Betas, and their significance. The OLS models were checked by tolerance tests to be free of any multi-collinearity problems. See appendix A for details about the measurement, coding, and construction of all variables. Significant coefficients are highlighted in **bold**.

**Source:** World Values Survey 2005-7.

---

<sup>1</sup> Robert Bickel. 2007. *Multilevel Analysis for Applied Research: Its Just Regression!* New York: The Guilford Press.

<sup>2</sup> Stephen W. Raudenbush and Anthony S. Bryk. 2002. *Hierarchical Linear Models* (2<sup>nd</sup> ed). Thousand Oaks: Sage;  
Andrew Gelman and Jennifer Hill. 2007. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. New York: Cambridge University Press.