[Appendices for Schakel]

<div align="center">

Appendix A

Expert Survey[11]

</div>

The functional variables used in the analysis in this article are based upon an expert survey conducted in January–March 2006 by Liesbet Hooghe, Gary Marks and Arjan H. Schakel. The experts were obtained from the member list of the Organized Section Federalism and Intergovernmental Relations of the American Political Science Association and of the European Group of Public Administration. The section and EGPA organizes members with an interest in federalism, intergovernmental relations and state and local government. Thirty–six out of 120 experts (response rate 30%) were asked to evaluate the externalities and scale effects for 34 policies. All experts were academics at American (30) or European (6) universities.

The 34 policies were taken from the country studies performed by the Council of Europe (1996–2000) and the Local Government and Public Reform Initiative (Horváth 2000; Kandeva 2001; Munteanu and Popa 2001) to establish congruence with the country policy provision dataset (see appendix B). The question wording for each of the 34 policies was as follows:

> "Please place yourself in the role of a public policy analyst and put an X in the boxes [jurisdiction] that best fit your evaluation of what levels of government are most efficient in providing the policy in question. We would like you to give your judgment abstracting from the particulars of any country (i.e. whether a policy is actually provided in a particular country). Also the question of which level of government funds the policy is a separate topic and should not affect your judgments in this survey. By efficiency, we refer to 1) economic externalities and 2) scale economies."

The question was followed by a definition of scale effects and economic externalities.

> "**Economic externalities** are the positive or negative economic effects of a policy for individuals in other jurisdictions. Efficient policy should encompass the people economically affected by the policy. For example, defense policy protects all those who live in a country, while street cleaning affects only those in a particular locality.

> "**Scale economies** refer to the decreased cost of policy provision per unit as the scale of provision increases. Efficient policy should reap the available economies of scale for providing a policy. Defense policy is most efficient when a single army deters threats to all those who live in a country, while street cleaning can be efficiently organized at a local level."

The expert was allowed to put an *X* in five jurisdictions (boxes) with assigned population sizes (based upon average population sizes of the jurisdictions classified in the Nomenclature of Territorial Units for Statistics (NUTS) and Local Administrative Unit (LAU) used by the European Union):

Local	< 20,000
Local–Regional	± 100,000
Regional	± 1 million
Regional–National	± 5 million
National	> 10 million

The experts were allowed to put an *X* in more than one jurisdiction to allow for the possibility that some policies are efficiently handled at multiple scales. A list of policies is provided in table A1. This table also provides the scores for the multilevel variable used in the analysis.

TABLE A1
List of 34 Policies Used in the Expert Survey

Policy	Multilevel
Roads (including local roads to highways)	101
Transport (including rail transport, subways/metro, buses)	99
Environmental protection (including air, water, soil)	95
Health protection (e.g. disease prevention)	86
Water supply	83
Tourism promotion	83
Museums	82
Sewage and water treatment	79
Public housing	78
Family welfare services (e.g. homeless shelters/families in crisis)	77
Secondary education	77
Parks and open spaces	76
Primary education	73
Vocational and technical education	73
Electricity	71
Welfare homes (e.g. orphanages)	71
Sport and leisure facilities	71
Libraries	70
In–home services for the elderly and the handicapped	70
Voter registration	69
Theatre and concert facilities	69
Hospitals	69
Gas	68
Regional/spatial planning	68
Town planning	68
Refuse collection	66
Refuse disposal	65
Higher education	64
Pre–school education	62
Consumer protection	61
Nursery and kindergarten	59
Fire protection	58
District heating (public distribution of hot water)	53
Cemeteries and crematoria	47

Notes: *Multilevel* is measured by the total number of placed *X*'s by all experts for a given policy; high values indicate that the policy should be provided by multiple governmental tiers when only externalities and scale effects are considered; low values indicate that a jurisdiction of a particular population size should provide the policy.

Inter–Expert Reliability

We use the Cronbach's alpha[12] for each jurisdiction and all jurisdictions combined to measure reliability among expert evaluations. From table A2 one may conclude that the experts converge.

TABLE A2
Cronbach's Alpha Scores Per Jurisdiction

	Local	Local–Regional	Regional	Regional–National	National	Overall
Cronbach's a	0.890	0.743	0.843	0.834	0.913	0.872

Notes: results for 34 experts (two experts had to be excluded as they had too many missing data) over 34 policies (listwise deletion).

Structural error

Two different versions of the expert survey were sent to the experts. Version A presented the 34 policies in the order as they appeared in the country studies mentioned above (N = 14). Version B presented the 34 policies in alphabetical order (N =22). The presence of systematic error due to the presentation of the list of policies can be gauged by comparing the answers of the experts for both versions of the expert survey. An one–way ANOVA analysis for each policy reveals that 9 out of 170 possible comparisons (34 policies x 5 jurisdictions) are significantly different between the two versions (that is 5%). This leads to the conclusion that, overall, there are no significant differences between the two versions of the expert survey and that there is no systematic error due to the listing of the policies.

Another structural bias may result from the fact that most consulted experts work at a university in the United States of America (the country, however, is not included in the analysis). It might be that their country experience (partly) framed their answers to the question which jurisdiction should provide a policy. We cannot discern whether this is the case but we may argue to what extent this has implications for the findings. The results are based upon differences between policy provision by tiers as functional theory would have it against actual policy provision in countries. If the experts used their country experience in their answers than the benchmark is biased and does not reflect optimal policy provision according to scale effects and externalities. Rather, the deviations refer to a difference in policy provision between the United States and another country. In spite of this, the conclusions remain the same; that is, when, for example, ethnic fragmentation rises it leads to a higher probability that the regional tier is involved in policy provision compared to the state level in the United States. The results, however, are either under– or overestimations compared to a functional benchmark. As the United States is a rather decentralized country, compared to the countries in our analysis, underestimation is more likely than overestimation.

Funding of Policies

It is important to note that the experts were asked to discard the question which level of government should fund the policies (see question wording above). Spillovers benefits across jurisdictions arising from local policy provision "can be promoted by appropriate unit subsidies which might encourage decentralized authorities to extend outputs to efficient levels" (Oates 2005, 352). Internalizing externalities constitutes, therefore, not a necessary condition for optimal policy provision. The exclusion of the question which level of government should finance the policy or which level of government should have which tax powers (Oates 2005) does not need to concerns us as I am interested in explaining local and regional policy provision and not whether and how optimal levels of policy output are being reached.

Validation

The expert survey on externalities and scale effects of 34 policies is, to our knowledge, the first one of its kind. Therefore, it is difficult to validate the data. Alesina, Angeloni amd Schuknecht (2005) compare their normative assessment of the desirable allocation of policy responsibilities between the European Union and member states with public opinion data (Eurobarometer). Unfortunately, public opinion data on the desired allocation of policy provision tasks between national and subnational tiers for the countries included in this analysis do not exist. To tentatively validate the data I use Shah's (1997) assessment on assigning responsibilities for local public services. Anwar Shah is an economist, coordinator of the public sector reform cluster at the World Bank and has advised many governments on fiscal federalism issues. For several policies he provides "a subjective assessment of how various allocative criteria favor either local or metropolitan assignment" (1997: 21). Shah uses economies of scale, economies of scope, benefit–cost spillout, political proximity and consumer sovereignty as allocative criteria which (partly) overlap with the criteria used by the experts during their assessment. Based on his evaluation Shah comes up with a categorization of policies: local public services that could be decentralization to (1) all local governments, (2) to larger urban municipalities and (3) to metropolitan or regional governments (1997: 22–24). What makes Shah's analysis informative is that he gives approximate population sizes for larger urban municipalities and metropolitan or regional governments which coincide with the population size assigned to the jurisdictions Local–Regional (approx. 100,000 people) and Regional (approx. 1 million people) in the expert survey.

   In table A3 we compare Shah's analysis with the expert survey for 19 policies. The first two columns give the policy labels as used by Shah and the expert survey. The next columns represent the proportion of experts which placed an X in that jurisdictional category (each expert is given a weight of one). Figures in bold indicate that a majority of experts agree with Shah whereas underlined figures indicate that a majority of experts places an X in another jurisdiction.

   As can be seen in table A3, the experts and Shah agree on the first five policies. A majority of the experts place an X in the local jurisdiction (< 20,000) for policies which, according to Shah, should be decentralized to all local governments. This is also the case for policies which could be decentralized to larger urban municipalities (population over 100,000). Most experts place an X in the jurisdiction Local–Regional (100,000) except for land use planning. However, in the latter case there is a difference in label use between the expert survey and Shah which might indicate that they refer to different policies. Shah makes a distinction between "land use planning" and "regional planning" whereas the experts only assessed "regional and spatial planning."

   There is also convergence between the experts and Shah for policies which, according to Shah, could be provided by metropolitan or regional governments (combined population of one million). Again, a difference in label use might explain disagreements. Shah use the labels "neighborhood parks and recreation," "regional parks," "local libraries" and "special libraries" whereas the experts were asked to evaluate "parks and open spaces" and "libraries" without further differentiation. Therefore, it is not surprising that the experts are somewhat evenly divided over the three smallest jurisdictions (i.e. Local, Local–Regional and Regional) for these two policies. The difference in label use might also explain the disagreement between Shah and the experts for "public health"/"health protection." The experts and Shah also diverge in their judgment for "hospitals" and "air and water pollution"/"environmental protection," though a second largest majority of the experts placed an X in a jurisdiction next to the preferred population size by Shah. Overall, Shah and the experts agree to a large extent and this we find comforting.

TABLE A3
Comparison between Shah (1997) and the Expert Survey on the Question, Which Jurisdiction Should Provide the Policy, for 19 Policies

Policy label used by Shah (1997)	Policy label used in the expert survey	Jurisdictional sizes as used in the expert survey				
		Local < 20,000	Local−Regional ± 100,000	Regional ± 1 million	Regional−National ± 5 million	National > 10 million
Local public services that could be decentralized to all local governments according to Shah (1997)						
Fire protection	Fire protection	**0.52**	0.27	0.14	0.01	0.06
Primary education	Primary education	0.34	<u>0.36</u>	0.14	0.09	0.07
Refuse collection	Refuse collection	**0.45**	0.26	0.16	0.08	0.05
Neighborhood parks and recreation	Parks and open spaces	**0.39**	0.26	0.25	0.06	0.04
Local libraries	Libraries	**0.29**	0.29	0.28	0.07	0.06
Local public services that could be decentralized to larger urban municipalities (population over 100,000) according to Shah (1997)						
Land use planning	Regional/spatial planning	0.05	0.24	<u>0.40</u>	0.24	0.07
Secondary education	Secondary education	0.21	**0.37**	0.23	0.07	0.12
Local public services to be provided by metropolitan or regional governments (combined population of one million) according to Shah (1997)						
Air and water pollution	Environmental protection	0.10	0.14	0.18	<u>0.32</u>	0.27
Electric power	Electricity	0.12	0.17	**0.37**	0.22	0.13
Gas	Gas	0.19	0.20	**0.31**	0.16	0.14
Hospitals	Hospitals	0.07	<u>0.39</u>	0.24	0.21	0.09
Public health	Health protection	0.09	0.16	0.19	0.22	<u>0.34</u>
Refuse disposal	Refuse disposal	0.23	0.24	**0.29**	0.16	0.08
Regional parks	Parks and open spaces	<u>0.39</u>	0.26	0.25	0.06	0.04
Regional planning.	Regional/spatial planning	0.05	0.24	**0.40**	0.24	0.07
Sewage disposal	Sewage and water treatment	0.19	0.29	**0.32**	0.13	0.07
Special libraries	Libraries	<u>0.29</u>	<u>0.29</u>	0.28	0.07	0.06
Transportation	Transport	0.08	0.19	**0.31**	0.26	0.16
Water supply	Water supply	0.22	0.22	**0.30**	0.17	0.09

Notes: The figures represent the proportion of experts which placed an X in that jurisdictional category (each expert is given a weight of 1).
Figures in bold represent agreement whereas underlined figures represent disagreement between Shah (1997) and a majority of the experts.

Appendix B
Country Data

The Council of Europe (CoE) has published 32 country studies on the structure and operation of local and regional democracy (Council of Europe 1996–2000). Twenty–seven country studies contain a table which shows which tiers are competent for the provision of 47 policies. Representatives from the ministries in charge of local and regional government represented the country in the CDLR (The Steering Committee on Local and Regional Democracy of the Council of Europe) and they completed/filled in the tables (*pers.comm.* Montgomery). These tables also indicate the type of competence, i.e. whether the tier of government has exclusive, shared, compulsory, or discretionary competence, and the exercise of the competence (direct, indirect, in own right, of for another authority). Since the definitions may not have been consistently applied across countries (*pers.comm.* Montgomery), I have chosen not use this information.

The Local Government and Public Service Reform Initiative (LGI) conducted country studies in Eastern and Central European Countries and in several former Russian Republics (Horváth 2000; Kandeva 2001; Munteanu and Popa 2001). The set–up of the country studies and the information contained within the country studies is broadly similar to that of the CoE. Twenty–three country studies present tables which score for 44 policies whether a governmental tier has a role in policy provision.

The country studies provide data on actual policy provision per tier for a total of 40 countries (16 West European, 15 Central and Eastern European and 9 Caucasian republics) and 34 policies common to both surveys (see table A1). Ten countries are analyzed by both sources (Bulgaria, Croatia, Czech Republic, Estonia, Latvia, Lithuania, Macedonia, Poland, Romania and Slovakia). To enhance comparability, the Council of Europe data for these ten countries are used since the CoE dataset is the source of most countries. The results presented in this article do not significantly change when data from the LGI country studies are used instead of the CoE data.

Three–Tier versus Two–Tier Countries

The decision whether to decentralize policy provision is highly dependent on the number of subnational tiers. When a policy is best provided by a regional tier but a country has no regional tier, the government has to decide whether to provide the policy at the national or local level or to provide it at both levels. However, in a three–tier country the policy can be provided at the regional level. This means that the possible *choices* regarding which tiers should provide a policy differ in three–tier versus two–tier countries. In order to account for this, the dataset is split up in two databases. One database refers 14 two–tier countries. These are: Armenia, Azerbaijan, Bulgaria, Cyprus, Czech Republic, Estonia, Iceland, Luxembourg, Macedonia, Malta, Portugal, Slovakia, Slovenia and Turkey.

The second database refers to 26 three–tier countries and these are: Albania, Austria, Belarus, Belgium, Bosnia and Herzegovina, Croatia, Denmark, Finland, France, Greece, Hungary, Kazakhstan, Latvia, Lithuania, Moldova, the Netherlands, Norway, Poland, Romania, Russian federation, Serbia and Montenegro, Spain, Sweden, Switzerland, Ukraine and Uzbekistan.

Five of these three–tier countries have actually four governmental tiers (i.e. Belgium, Bosnia and Herzegovina, France, Spain, and Poland). There are several reasons why these countries are not analyzed separately. First, in these countries there are fifteen possible combinations of governmental tiers which may provide a policy (i.e. N, R_1, R_2, L, NR_1, NR_2, NL, R_1R_2, R_1L, R_2L, NR_1R_2L, NR_1R_2, NR_1L, NR_2L, R_1R_2L). Some of the categories are empty—that is, policy provision is never provided by a specific combination—and multinomial models cannot be estimated when certain categories do not have observations.

Second, because policies are embedded in countries the number of "true" observations in a statistical sense is far closer to five than to 170 (5 countries times 34 policies) which makes the likelihood that one finds statistical significant results rather low. Furthermore, a low number of cases will give rise to estimation difficulties due to multicollinearity between the independent variables. For these reasons,

the four–tier countries are included in the three–tier country database by taking the data for the most authoritative intermediate tier while disregarding data of the less authoritative intermediate tier. The results appear to be robust when data for the less authoritative tier is included.

Selection Bias

The policies examined in the surveys are not randomly selected. They are policies where it is conceivable that subnational tiers may play a role. This is why the dataset does not contain policies that tend to be exclusively provided by the national tier, for example defense, foreign relations and justice. Hence one should be aware of a potential selection bias in the data, though it has to be said that a large number of selected policies involve the national government as policy provider: 42 percent of combinations in three–tier countries, and 75 percent in two–tier countries (see table B1). If there is a selection bias it is not likely to be severe.

TABLE B1
Frequency of Policy Provision by Government Tiers in Three– and Two–Tier Countries

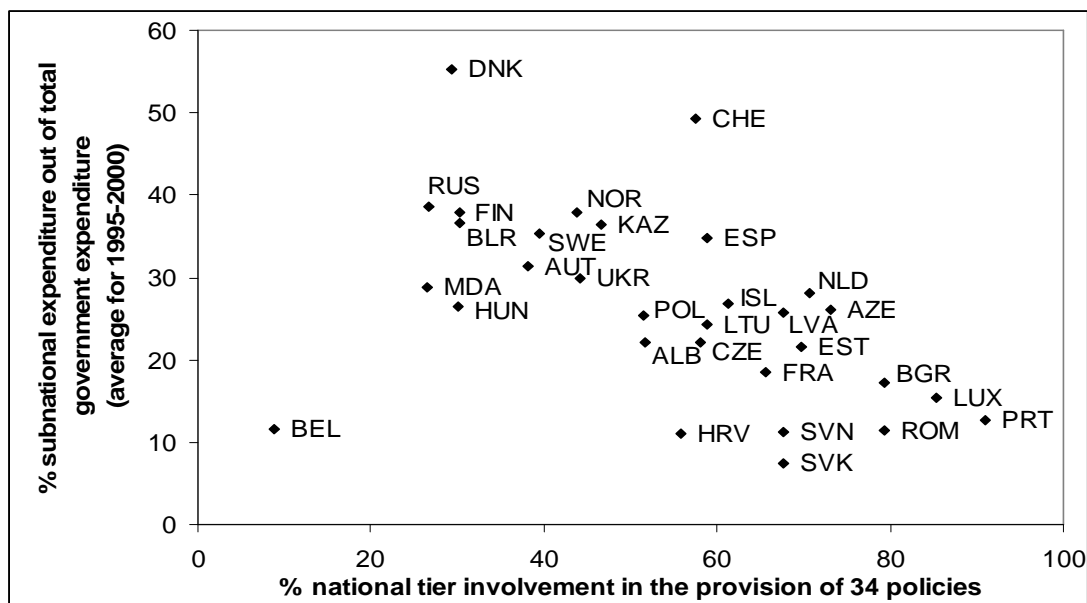|	Three-tier countries		Two-tier countries	
Policy provision by the	Frequency	Percentage	Frequency	Percentage
National tier only	60	7.0	114	24.9
Regional tier only	55	6.5	–	–
Local tier only	239	28.0	110	24.0
National+Regional tier	39	4.6	–	–
National+Local tier	84	9.9	234	51.1
Regional+Local tier	199	23.3	–	–
National+Regional+Local tier	177	20.8	–	–
Total	853	100%	458	100%

Notes: The frequency data refer to country*policies. Policy provision is measured for 34 policies in 26 three–tier and 14 two–tier countries and refers to one year between 1996–2001.

Validation

The country data are new in its kind. To my knowledge, there does not exist another quantifiable and broadly comparative list of policies provided by governmental tiers. Thorlakson (2003, 6–11) develops a numerical index of federal jurisdiction for six federations; it measures the proportion of policy areas of exclusive state jurisdiction compared to areas of concurrent, shared and exclusive federal jurisdiction. Alesina, Angeloni and Schuknecht (2005, 279) measure the policy involvement of the European Union—but not the member states—by counting "the number of legal, judiciary and other, non–binding acts emanating from the European Union". Most often, scholars use subnational expenditure data as a proxy for subnational policy involvement—despite the caveats associated with these indicators (see Schakel 2008). In order to gauge the validity of the dependent variable I compare the dependent variable with subnational expenditure data. To this purpose I need to re–operationalize the country data. I follow the approach taken by Alesina, Angeloni and Schuknecht (2005) and count the number of policies for which the national tier is responsible (i.e. N, NR, NL and NRL). Subsequently, I calculate the proportion of policies for which the national tier is involved out of the total number of policies. Figure B1 plots this measure against the percentage of subnational expenditure out of total government expenditure.

FIGURE B1

National Tier Involvement in the Provision of 34 Policies (in Percentage) Plotted against the Percentage of Subnational Expenditure out of Total Government Expenditure (Average for 1995–2000) for 31 Countries

Notes: ALB = Albania; AUT = Austria; AZE = Azerbaijan; BEL = Belgium; BGR = Bulgaria; BLR = Belarus; CHE = Switzerland; CZE = Czech Republic; DNK = Denmark; ESP = Spain; EST = Estonia; FIN = Finland; FRA = France; HRV = Croatia; HUN = Hungary; ISL = Iceland; KAZ = Kazakhstan; LTU = Lithuania; LVA = Latvia; LUX = Luxembourg; MDA = Moldova; NLD = the Netherlands; NOR = Norway; POL = Poland; PRT = Portugal; ROM = Romania; RUS = Russian federation; SVK = Slovakia; SVN = Slovenia; SWE = Sweden; UKR = Ukraine.
Source: World Bank (2006)

As expected, national tier involvement in policies is negatively correlated with subnational expenditure (Pearson r: –0.50, $p < 0.01$, N = 31). Belgium is a clear outlier. This can be explained by a difference in timing of the policy decentralization and fiscal decentralization. The constitutional revision of 1993, which devolved competencies to the regions and communities, came into effect in 1995 but fiscal arrangements were only significantly revised in favor of regions and communities in 2001 (Swenden 2006). The measure for policy provision dates from 1997 and so falls between the two reforms. If Belgium is excluded the Pearson correlation is: –0.67 ($p < 0.01$, N = 30).

Appendix C
Descriptive Tables

TABLE C1
Descriptive Statistics of the Independent Variables for the Three–Tier and Two–Tier Country Dataset

Independent Variables	Three–tier countries			Two–tier countries		
	Mean	Std. deviation	Min – Max	Mean	Std. deviation	Min – Max
Percentage National tier	0.135	0.107	0.01 – 0.61	0.187	0.115	0.01 – 0.61
Percentage Regional tier	0.239	0.095	0.01 – 0.53	–	–	–
Percentage Local tier	0.251	0.145	0.02 – 0.53	0.251	0.151	0.03 – 0.53
Ceiling effect	0.057	0.113	0.00 – 0.81	0.198	0.230	0.00 – 0.93
Multilevel	72.50	11.55	47 – 101	72.82	11.21	47 – 101
Ethnic fragmentation	0.335	0.202	0.047 – 0.681	0.261	0.152	0.03 – 0.535
Strength ethnic/regional party	3.250	8.111	0.00 – 40.48	3.454	6.079	0.00 – 20.00
Democratic openness	27.483	10.818	0.19 – 42.03	29.746	9.345	6.87 – 43.62
Polyarchy	5.661	6.298	–9 – 10	7.633	4.009	–7 – 10
Economic welfare	9.174	0.815	7.659 – 10.177	9.225	0.696	8.055 – 10.514
EU–membership	0.239	0.425	0 – 1	0.146	0.354	0 – 1
EU–subsidies	0.062	0.220	0.00 – 1.078	0.111	0.382	0.00 – 1.480

Notes: data refer to 853 country*policies (26 countries) for three–tier countries and to 458 country*policies (14 countries) for two–tier countries.

TABLE C2
Pearson Correlations between the Independent Variables for the Three–Tier And Two–Tier Country Dataset

	Independent Variables	1	2	3	4	5	6	7	8	9	10	11	12
1	Percentage National tier	–	–	–0.36*	0.34*	–0.05	0.08	–0.04	0.03	0.20*	0.22*	0.00	–0.15
2	Percentage Regional tier	–0.23*	–	–	–	–	–	–	–	–	–	–	–
3	Percentage Local tier	–0.49*	–0.09*	–	–0.38*	0.02	–0.00	–0.01	–0.04	–0.05	–0.01	–0.01	–0.01
4	Ceiling effect	0.25*	–0.21*	–0.25*	–	–0.05	–0.08	–0.16*	0.14*	0.29*	0.28*	0.10*	–0.23*
5	Multilevel	0.28*	0.08*	–0.54*	0.11*	–	–0.01	0.01	0.03	0.03	0.02	–0.00	–0.00
6	Ethnic fragmentation	0.07	–0.13*	0.01	0.10*	–0.00	–	0.57*	–0.15*	–0.09	0.09*	–0.24*	–0.40*
7	Strength of ethnoregional parties	–0.02	–0.07*	0.00	0.06	0.00	0.40*	–	–0.01	0.03	–0.00	–0.24*	–0.17*
8	Democratic openness	0.02	–0.03	–0.01	–0.03	–0.00	–0.39*	0.11*	–	0.80*	0.53*	0.06	0.00
9	Polyarchy	0.04	0.08*	–0.01	0.11*	–0.00	–0.47*	–0.11*	0.78*	–	0.43*	0.24*	0.17*
10	Economic welfare	–0.05	0.04	–0.01	–0.08	–0.01	–0.42*	–0.25*	0.59*	0.56*	–	0.29*	0.13*
11	EU–membership	–0.06	–0.06	–0.01	–0.18*	–0.01	–0.19*	0.01	0.53*	0.37*	0.50*	–	0.70*
12	EU–subsidies	–0.04	0.02	–0.00	–0.14*	–0.00	–0.21*	–0.03	0.24*	0.19*	0.15*	0.38*	–

Notes: Pearson correlations for the three–tier country dataset (N = 853) are shown below the diagonal and the Pearson correlations for the two–tier country dataset (N = 453) are shown above the diagonal.
* $p < 0.05$