Julian L. Garritzmann, Leonce Röth, and Hanna Kleider:

Policy-making in Multi-Level Systems:

Ideology, Authority, and Education

Comparative Political Studies

Online Appendix I: Data, Sources, Descriptive Statistics, & Robustness

1. Data and Sources

Table A1: Data and sources

Variable	Obs.	Mean	Std. Dev.	Min	Max	Description and source
Dependent Variable						
Regional education spending per capita in int. $	3862	929.01	647.04	0.00	6163.08	Per capita spending, including public primary, secondary and higher education expenses. Own compilation based on national statistic offices and the COFOG classification scheme. Deflated and transformed in international $ to the base of 2010.
Main Independent Variables						
Regional authority over education (REAI)	6048	2.52	2.50	0.00	6.00	Multiplicative Index of educational scope and educational depth on the regional level (see below for a detailed description).
Depth of regional education	6048	1.65	1.05	0.00	3.00	See Part 2 of this Online Appendix.
Scope of regional education	6048	1.05	0.82	0.00	2.00	See Part 2 of this Online Appendix.
Ideology of Regional Government	6070	0.58	0.15	0.02	0.92	Regional government ideology is a cabinet seat share weighted measure cabinet parties' position. The position is based on the MARPOR (Volkens et al. 2016) data using the transformation procedure of (Röth 2017) for the economic state-market dimension.
Institutions						
Regional fiscal authority	6148	2.92	1.00	0.00	4.00	Fiscal authority captures the autonomy of regions to gain revenues from their own taxes. We rely on the coding of the sub dimension of the Regional Authority Index (Hooghe et al. 2016).
Regional Authority Index	6233	17.97	6.39	0.00	27.00	Regional Authority captures the general degree of authority that regions have. We rely on the coding of the Regional Authority Index (Hooghe et al. 2016).
RAI self-rule	6233	13.81	2.89	0.00	18.00	The self-rule dimension of the Regional Authority Index captures the autonomous competencies of regions in contrast to shared-rule or co-determination rights of regions in order to affect state-wide political decisions. We rely on the coding of the Regional Authority Index (Hooghe et al. 2016).
Multilingual Region	6233	0.12	0.323	0.00	1.00	Region with multilingualism in education are binary coded. We code regions as being multilingual when at least two languages have equal status of being taught as the primary language at school. See also Online Appendix Part 4.

Variable	Obs.	Mean	Std. Dev.	Min	Max	Description and source
Number of regions	6233	28.82	14.63	4.00	50.00	Number of regions per country and year. We use the number of regions covered by the data. Accordingly, uncovered but existing regions are not counted (own compilation).
Political Factors						
State-wide government ideology	6091	0.58	0.13	0.14	0.92	State-wide Government Ideology is a cabinet-seat-share weighted measure of the cabinets' party position on the state-market dimension. The state-market dimension is measured by the approach of Röth (2017).
Ideological proximity	5963	0.87	0.12	0.17	1.00	Ideological proximity is the ideological distance between the regional government and the national government. Government ideology on both levels include coalition governments. Parties in coalitions are weighted by cabinet seat shares. Own compilation.
Number of parties in reg. government	5839	1.64	1.32	1.00	10.00	Number of parties in regional government (own compilation).
Duration of government	6233	347.99	41.99	17.00	365.00	Duration of cabinet in days for respective year (own compilation).
Socio-Demographic Factors						
Mountainous and populated region	6118	0.70	0.84	0.00	2.00	Mountainous regions closed to the classification of Eurostat. (2018): 0 = majority of population lives in non-mountainous regions. 1 = regions with more than 50% of their population living in mountain areas; 2 = regions with more than 50% of their surface covered by mountain areas, and with more than 50% of their population living in mountain areas. The original classification is on the NUTS-3 Level and we adapted it to our territorial units of interest. Estimates for regions outside the EU are based on data of ARCGIS (2016) and population density data.
Percentage of population <15	5910	26.87	4.88	14.67	51.29	OECD (2016), compare Supplementary Material I.
Population density (log)	5915	4.32	1.79	-3.51	8.84	OECD (2016)), compare Supplementary Material I.
Regional GDP per capita in int. $ (in 1000)	5426	33.79	9.99	9.60	99.84	OECD (2016), compare Supplementary Material I.
Growth rate of regional GDP	5425	1.17	3.54	-21.59	40.80	OECD (2016), compare Supplementary Material I.
Rate of unemployment	4860	6.52	3.86	0.05	33.90	OECD (2016), compare Supplementary Material I.
Special regions	6233	0.05	0.21	0.00	1.00	Special region takes account of regions where we are unable to distinguish spending data regional and lower-level municipalities. The variable special region takes on a value of 1 if the nature of expenditure data does not allow the separation of regional and lower-level municipal functions, and a value of 0 in all other cases

Table A2: Overview REAI (for a detailed discussion of coding decisions see Online Appendix II)

Country	Region	Scope (general)	Scope (non-tertiary)	Scope (tertiary)	Depth (general)	Depth (non-tertiary)	Depth (tertiary)	REAI (general)
Australia	All regions	1	2	0	2	3	1.5	2
Austria	All regions	0 *de jure*; but 0.5 *de facto*	-	-	1	-	-	0 de jure, 0.5 de facto
Belgium (1)	French-, and Flemish-speaking communities and Brussels	2	2	2	2.5	2.5	2.5	5
Belgium (2)	German-speaking community	1	2	0	1.5	2.5	0	1.5
Canada	All regions	2	2	2	3	3	3	6
Denmark	All regions	0	-	-	0.5	-	-	0
France	All regions (except extra-territorial ones)	0.5	-	-	1	-	-	0.5
Germany	All regions	2	2	2	3	3	3	6
Italy (1)	All regions (except special status)	0	-	-	1	-	-	0
Italy (2)	Trentino-Alto Adige/Südtirol (later split in two provinces: Trentino/Province of Trento and South Tyrol/Province of Bolzano), and Valle d'Aosta	2	-	-	2	-	-	4
Italy (3)	Sicily	0.5	-	-	2 de jure 2, 1 de facto.	-	-	1 de jure, 0.5 de facto
Italy (4)	Sardinia and Friuli-Venezia Giulia	0	-	-	1	-	-	0
Japan	All regions	1	-	-	1	-	-	1
Norway	All regions	0	-	-	0.5	-	-	0
Spain	All regions	There is considerable change over time, see the details in the Online Appendix	There is considerable change over time, see the details in the Online Appendix	There is considerable change over time, see the details in the Online Appendix	There is considerable change over time, see the details in the Online Appendix	There is considerable change over time, see the details in the Online Appendix	There is considerable change over time, see the details in the Online Appendix	There is considerable change over time, see the details in the Online Appendix
Sweden	All regions	0	-	-	0	-	-	0
Switzerland	All regions	1.5	2	1	2.5	3	2	3.75
UK (1)	Scotland & Wales	0 until 1997, and 2 since 1998	-	-	0 until 1997, and 3 since 1998	-	-	0 until 1997, 6 since 1998
UK (2)	Northern Ireland	0 until 1999, 2 in 2000-2002, 0 in 2003-2006, and 2 since 2007	-	-	0 until 1999, 3 in 2000-2002, 0 in 2003-2006, and 3 since 2007	-	-	0 until 1999, 6 between 2000-02, 0 between 2003-06, 6 since 2007
UK (3)	England	0	-	-	0	-	-	0
USA	All regions	2	2	2	3	3	3	6

3. Regions with multiple language in education

Regions with multilingualism in education are binary coded. We code regions as being multilingual when at least two languages have equal status of being taught as the primary language at school. However, as our argument related to education spending, we do not code regions as multilingual where the size of a language minority is very small. For example, in northern Finland or Sweden around 470 people are educated in Sámi, 115 of them exclusively. We do not expect spending effects of such a small group of people and accordingly abstain from coding them multilingual. All coding decisions ae listed below.

Table A3: Coding of regions with multiple languages in education

Region	Country	Multilanguage	Selection Status
Kärnten	Austria	Slovenian minority and language is a salient political issue. However, not implemented in schooling.	Ambiguous
Vlanders, Wallonie, Brussels Capital	Belgium	Brussels is bilingual in education. Flanders and Wallonia are monolingual but specific Municipalities where the French/Dutch speaking minority has a right to be educated in mother tongue.	Multilingual regions
All provinces of Canada		French or English first policy entitles parents who have received English or French education to choose the same for their children.	Multilingual regions
Greenland	Denmark	For pupils with the Greenlandic language as their mother tongue, the subject of Danish may be introduced in the second grade, and has to be introduced by the fourth. For other pupils' special training in Greenlandic as a foreign language is given. Since Home Rule Act in 1979.	Multilingual regions
Friuli-Venezia Giulia, Piemonte, Trentino, region Valled'Aosta, Veneto, Südtirol	Italy	Various forms of regional multilingual policies on the regional level.	Multilingual regions
Alsace, Corse	France	In 1992 the French Ministry of Education finally agreed to start with a bilingual education system in Alsace leading to a steady increase of bilingual enrolment. Bilingual education in Corse started in 1996.	Multilingual regions (Alsace since 1992; Corse since 1996)
Cataluna, Navarra, Pays Basques, Galicia	Spain	Various forms of multilingual education policies in the different historical regions.	Multilingual regions
Jämtland, Norrbotten, Västerbotten, Västernorrland,	Sweden	Sámis are entitled to be educated in Sámi. Language policy in education is administered by the municipality level and only very few people are affected.	Not selected because the Sami population is too small and we do not expected substantial spending related effects.
Bern, Freiburg, Graubünden, Ticino, Wallis	Switzerland	Different forms of multilingual policies, partly harmonized by Swiss federal state law.	Multilingual regions

Note: Comparing average spending between multilingual regions is instructive for within countries comparison only because of the differences in education authority. Multilingual regions are usually situated in high sub-national authority contexts. In all cases the spending is higher in multilanguage regions in comparison to the non-multilingual regions. However, only in Italy are fundamental differences (average of 760 pc spending versus 209 in non-multilingual regions).

4. Robustness

The robustness section is structured into six parts. First, we analyse the importance of missing values. Second, we assess alternative model specification strategies and conclude that the models presented in the paper are the most appropriate as well as the most conservative estimates for the effects of main concern. Third, we put our index of regional authority over education under scrutiny and check if alternative aggregation procedures or alternative indices such as the regional authority index would have led to substantially different results of regional government ideology and policy making in education. Fourth, we assess whether specific countries are of substantial importance for our findings and discuss their importance once we find substantially altered coefficients. Fifth, we analyse the importance of specific controls in order to see if our findings depend on the inclusion of specific controls and discuss the upper and lower bound of our findings. Sixth, we compare whether other ideological scales such as the RILE would alter our findings. Finally, we analyse the conditionality of the ideology effects and interacts government ideology in decentralized regions with opportunity structures. To pre-empt the findings on the ladder, we see that majority, time or socio-economic conditions hardly matter whereas ideological proximity to the state-wide government is of fundamental importance.

4.1 Missings on the dependent variable

We distinguish between three types of missings and evaluate whether their characteristics give guidance to systematic and problematic exclusions of key observations. The first group of missing cases are those where we have data in principle but for specific region-years we are short of education spending data. The observations which fall into this groups are mainly from the USA and a few region-years in Denmark and Canada. We are unable to receive data for education of US states in 1990, 1991, 2001, 2003 and 2007. In total we lose 247 region-years from the US and another 15 from Canada and Denmark which obviously share mainly all the special characteristic of US states.

The second type of missings concerns regions where no spending data exist at any time point. Those are for example the Germanophone region in Belgium, the Faroe Islands, Corse or Ceuta and Melilla. In total they sum up to 248 region-years. Those cases have so few observations on many other controls that we cannot even estimate treatment and control balances or "missing at random" models. Full information are available on the institutional characteristics of the regional authority index. They rank only slightly below the average of the entire sample (self-rule 13 versus 13.85 and shared rule 3.01 versus 4.21).

The third and by far most numerous group of missing cases are those where observations on education spending is not available for longer time periods. In total those are 1,876 region years and consist of all Japanese regions between 1990 and 2000, the UK regions between 1990 and 2000, Swedish regions between 1990 and 2001, the Spanish regions before they received authority over education (different years across regions), Norwegian regions between 1999 and 2010, Italian regions between 1990 and 2000; German regions between 1990 and 1994, Belgian regions between 1990 and 2000, Northwest Territories in Canada between 1990 and 1999, the Danish Amter between 1990 and 1994 and the regions in France between 1990 and 2000. A closer look at the distribution indicates that missing are not random. Missing regions are less likely to have multilingual arrangements, they have lower

populations, less authority over education and in lower authority in general (self-rule and shared-rule). Due to the high number of observations, even marginal differences such as the likelihood to be a mountainous region are significant (see results from the Probit model comparing the covariates of missing and non-missing variables). Most importantly, the main treatment variable (regional government ideology) is not substantially different. However, the substantially lower regional education authority in the missing cases is more problematic. We know that in countries like Spain, regional authority over education and the existence of regional spending data coincide. We take that as evidence that in "centralized" settings regional governments can hardly influence education spending when respective regional budgetary positions are not even existent.

Overall, we abstain from any imputation of values for the dependent variable in our analysis. Furthermore, we point to the fact that our sample is biased towards cases with higher regional authority which in tendency coincides with data availability. Since we assume and find null-findings for contexts of low regional authority and strong findings for high regional authority contexts, we do not see problematic implications for cautious generalizations of our findings.

Table A4: Distribution of covariates across missing cases and non-missing cases

	Mean (missings)	Mean (non-missings	Coefficient probit model	P>|z|
Regional cabinet ideology	0.58	0.58	0.41	0.007
Multilingual education	0.04	0.13	-0.42	0.000
Mountain region	0.69	0.72	-0.09	0.002
Log population density	4.60	4.21	0.07	0.000
Percentage of population under 14	26.43	27.00	0.03	0.000
Special region	0.06	0.04	0.35	0.001
Population	20.87	30.97	-0.01	0.000
Number of region per country/year	29.43	29.90	0.05	0.000
Duration of government in days/year	346.1	348.3	-0.00	0.038
Seat share government	0.68	0.78	-1.61	0.000
Nr of parties in government	1.49	1.69	0.16	0.000
Ideology state-wide government	0.55	0.59	-0.62	0.000
RAI (self-rule dimension)	11.71	14.83	-0.45	0.000
RAI (shared-rule dimension)	1.41	5.40	0.14	0.000
Regional authority over education	0.80	3.28	-0.18	0.000

N= 5696

4.2 Alternative Specifications

In this section we assess whether alternative specifications would have led to contradictory evidence and inference. In the paper we argue for a three-level multilevel model because we see education policy, although driven by regional governments, nested in state-wide and regional structures. To declare education policy across regional governments completely independent would thus violate our theoretical assumptions and run counter to the knowledge of education policy-making in the field. Nonetheless, statistical concerns might be brought forward because the number of countries (15) violates several rule-of-thumb suggestions for multi-level modelling. For that reason, we alternatively estimate two-level models with regional governments being nested in regions only. The coefficients of the interaction between regional government ideology and the level of regional education authority is very similar (compare model 1 and 2, Table B1). This finding does not vary once we check for random versus fixed effects within the two-level framework (compare model 1,2 and 3).

Adding a lagged dependent variable does also not change the effect of main concern substantially and thus provides leverage to infer that endogeneity of education spending and regional government ideology is widespread (see model 4). Leaving the multi-level framework and treat regional governments as purely independent leads to substantially stronger effects (model 5 and 6) irrespective of a fixed or random-effects specification. We interpret these effects as over-estimated. Once we deliberately and additionally abstain from temporal dependencies and run a simple OLS regression, we would arrive at an effect which is around 9 times the effect we present in the main text (compare model 7). These findings motivate our conclusion that the presented effects are conservative estimates and ignoring temporal and spatial dependencies would have had led to a massive increase in effect sizes, an increase we describe as over-estimation violating our theoretical assumptions on regional policy-making in education.

Table A5: *Robustness– Alternative Specifications*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Estimator	ML with random effects	ML with random effects	ML with fixed effects	ML with fixed effects	Panel fixed effects	Panel random effects	OLS
DV: Education spending per capita	Primary, secondary, tertiary	Primary, secondary, tertiary	Primary, secondary, tertiary	Primary, secondary, tertiary	Primary, secondary, tertiary	Primary, secondary, tertiary	Primary, secondary, tertiary
Nested level	Country, Region	Region	Region	Region	-	-	-
REAI Indicator	Scope*Depth	Scope*Depth	Scope*Depth	Scope*Depth	Scope*Depth	Scope*Depth	Scope*Depth
Lagged DV	no	no	no	yes	no	no	no
Regional cabinet ideology	0.62	18.54	9.81	11.48	14.33	27.25	409.12***
State-wide cabinet ideology	-99.67***	-96.18***	-57.71***	-6.50	18.34	16.81	453.65***
Ideological proximity	36.59*	40.76**	26.02	-18.95	-11.83	-9.47	422.31***
Regional cabinet ideology * Regional authority over education	-23.18***	-23.95***	-22.47***	-21.18***	-32.73***	-40.89***	-182.31***
Regional authority over education	23.21**	54.86***	13.02	14.68	17.93*	76.43***	239.49***
Std. of regional gov. ideology	.15	.15	.15	.15	.15	.15	.15
Std. of state-wide gov. ideology	.13	.13	.13	.13	.13	.13	.13
n	3739	3739	3739	3326	3739	3739	3739
countries	15	15	15	15	15	15	15
regions	282	282	282	282	282	282	282

Controls: Duration reg. cabinet in days, **n**umber of parties in cabinet, Percentage of population < 14, Regional population (in 100.000), log of population density, regional GDP per capita (in 1000), Regional GDP growth (in %), Mountain region, Multilingual education, Unemployment, Special region, Nr. of regions per country.

4.3 Alternative Aggregations and Authority Indices

An important element of our theoretical framework is the degree of regional authority over education policy. While Hooghe et al. (2016) recently proposed a measure of regional authority, their measure is not policy-specific. As we have strong reasons to assume that education is special when it comes to authority, using a more general measures of regional governments' authority to proxy for competencies in education policy might be misleading. Thus, we created a novel measure of regional authority over education policy, as explained above. In developing this measure, we used the established procedure by Hooghe et al. (2016) as our starting point and distinguish between a 'scope' and 'depth' dimension.[1] We combine the scope and the depth dimension into a more general index of regional governments' authority over education policy. We connect both dimensions multiplicatively, because both dimensions are necessary but not individually sufficient to exercise influence over education policy. To exemplify, if a regional government has some authority over all education areas, but this authority is strictly limited by the central government, then this competence does not mean a lot. Thus, it seems plausible to weight the scope of their authority with their depth. Therefore we create the following Regional Education Authority Index (REAI) that takes values for each region *r* at time *t*:

$$Regional\ Education\ Authority\ Index_{rt} = Scope_{rt} \times Depth_{rt} \qquad (1)$$

In this section, we present some alternative specifications and test the implications for our core hypotheses 1. First, we test the effect of the individual components, namely, scope and depth separately. Second, we also assess additive combinations for those who perceive scope and depth as functional equivalents. Third, we add fiscal authority (borrowed from the constituent dimensions of the RAI) multiplicatively and additively to scope and depth. Finally, we test the RAI and its self-rule dimension as alternative ways of mapping regional authority over education. Table B2 summarizes the correlations between the different indices. Regions are typically equipped with scope and depth leading to a correlation of 0.92. Accordingly, using indices for simply scope or depth results in comparable effects to its additive connection (compare results in Table B2 and Table B2). Although empirical differences between multiplicative and additively developed indices are not very pronounced in our sample, we think the multiplicative combination of scope and depth is the theoretically the most compelling. The distribution of fiscal authority is more loosely connected to scope and depth (correlation of 0.45 and 0.44). A note of caution is worth making here, effect sizes cannot be substantially compared across models, because the different indices have different scales. Adding fiscal authority multiplicatively is leading to weaker ideology effects of governments (in terms of significance) whereas adding fiscal authority additively is strengthening it. We can carefully infer, that fiscal authority is not necessary for regional governments to make a difference on education spending but it helps and adds to the competences captured by scope and depth (see model 8 and 9 in Table B3).

[1] Following Hooghe et al. (2016) we code regions' *de jure* powers. Yet, based on the scholarly literature and experts' evaluations, we also paid attention to circumstances where *de facto* powers differ. We explain these cases in the Codebook in the Online Appendix. In the empirical analysis we concentrate on the *de facto* powers.

Table A6: *Correlation of alternative authority scales*

	scope	**depth**
depth	0.92	1.00
Scope * Depth	0.95	0.96
Scope + Depth	0.97	0.98
Scope * Depth * Fiscal	0.87	0.91
Scope + Depth + Fiscal	0.91	0.91
Fiscal Authority	0.45	0.44
Regional Authority Index	0.71	0.84
Self-Rule Dimension (RAI)	0.63	0.76

Our pivotal hypothesis 1 is also confirmed once we replace our measure of regional education authority with the general index of regional authority (model 10). We thus conclude that the regional authority index captures competences over education roughly well. However, a closer look reveals that for example within-country differences over time and across regions in the authority over education are not well captured by the general regional authority index. For example, the correlation for Spain is 0.50 and even lower for Italy (0.32). Furthermore, even those modest correlations are based on aggregation effects of shared and self-rule. A conceptually proper comparison would be the correlation between the regional education authority index and the self-rule dimension of the RAI because our measure captures self-rule dimensions only. Looking at those correlation within countries reveal even bigger differences between both measures. The correlation of the self-rule dimension in Spain is, for example, 0.30 (In Italy 0.21; in Belgium 0.38; in Australia 0.26). The self-rule dimension of the RAI captures differences in the REAI fairly well only in a single country (the UK with correlation of 0.95).

Table A7: *Robustness– Alternative REAI Indicators*

	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Estimator	ML with region random effects	ML with random effects	ML with fixed effects	ML with random effects	ML with random effects	ML with random effects	ML with random effects
DV: Education spending per capita	Primary, secondary, tertiary	Primary, secondary, tertiary	Primary, secondary, tertiary	Primary, secondary, tertiary	Primary, secondary, tertiary	Primary, secondary, tertiary	Primary, secondary, tertiary
Nested level	Country, Region	Region	Region	Region	Region	Region	Region
REAI Indicator	Scope*Depth*Fiscal	Scope + Depth + Fiscal	Regional Authority Index	Self-Rule	Scope	Depth	Scope + Depth
Regional cabinet ideology	-0.58	34.38	152.38**	159.15	34.38	69.65	57.79
State-wide cabinet ideology	-97.79***	-96.06***	-91.08***	-94.70***	-96.06***	-95.21***	-95.63***
Ideological proximity	38.10*	41.16**	41.06**	38.72*	41.16**	41.32**	41.47**
Regional cabinet ideology * Regional authority over education	-5.21**	-77.88***	-10.75***	-15.08*	-77.88***	-66.17***	-37.15***
Regional authority over education	21.06***	152.33***	25.40***	16.95***	152.33***	138.06***	77.83***
Std. of regional gov. ideology	.15	.15	.15	.15	.15	.15	.15
Std. of state-wide gov. ideology	.13	.13	.13	.13	.13	.13	.13
n	3739	3739	3739	3739	3739	3739	3739
countries	15	15	15	15	15	15	15
regions	282	282	282	282	282	282	282

Controls: Duration reg. cabinet in days, number of parties in cabinet, Percentage of population < 14, Regional population (in 100.000), Log of population density, Regional GDP per capita (in 1000), Regional GDP growth (in %), Mountain region, Multilingual education, Unemployment, Special region, Nr. of regions per country.

4.4 Country dependency

The evidence of this study is based on regions within 15 countries. Although, we control for the number of regions per country in order to avoid stronger impact of countries with more regions, single countries can be very influential for our inference in particular because our main argument is based on an interaction where regions in specific countries stand representative for a specific authority distribution. Hence, we case-wise exclude countries in order to see if single countries drive our results. None of the results indicate that a particular country is necessary to arrive at our core inference as the interaction effects remain all negative and highly significant. We could learn from the relative change of the slopes that for example the inclusion of Austria and Sweden strengthen our results whereas countries such as Canada or Australia slightly reduce the negative relationship between regional authority over education and the ideology of regional governments.

Table A8: *Robustness– Exclusion of Single Countries I*

	(1)	(2)	(3)	(4)	(5)	(6)
Estimator	ML with region fixed effects	ML with region fixed effects	ML with region fixed effects	ML with region fixed effects	ML with region fixed effects	ML with region fixed effects
DV: Education spending per capita	Primary, secondary, tertiary	Primary, secondary, tertiary	Primary, secondary, tertiary	Primary, secondary, tertiary	Primary, secondary, tertiary	Primary, secondary, tertiary
Nested level	Region	Region	Region	Region	Region	Region
REAI Indicator	Scope*Depth	Scope*Depth	Scope*Depth	Scope*Depth	Scope*Depth	Scope*Depth
Regional cabinet ideology	57.64**	-45.69	13.86	15.92	9.92	23.88
State-wide cabinet ideology	-41.44**	14.95	18.96	8.38	30.92	8.75
Ideological proximity	0.50	-13.21	-12.01	-24.97	-5.92	-7.50
Regional cabinet ideology * Regional authority over education	**-38.99*** **	**-21.43*** **	**-32.71*** **	**-43.04*** **	**-32.06*** **	**-34.32*** **
Regional authority over education	18.80*	12.02	17.93	18.33*	17.04*	18.95*
Std. of regional gov. ideology	.15	.15	.15	.15	.15	.15
Std. of state-wide gov. ideology	.13	.13	.13	.13	.13	.13
n	3571	3550	3712	3522	3576	3533
countries	14	14	14	14	14	14
Country Excluded	Australia	Austria	Belgium	Canada	Denmark	France
regions	274	273	279	271	268	261

Controls: Duration reg. cabinet in days, number of parties in cabinet, Percentage of population < 14, Regional population (in 100.000), Log of population density, Regional GDP per capita (in 1000), Regional GDP growth (in %), Mountain region, Multilingual education, Unemployment, Special region, Nr. of regions per country.

Table A9: *Robustness– Exclusion of Single Countries II*

	(7)	(8)	(9)	(10)	(11)	(12)
Estimator	ML with region fixed effects	ML with region fixed effects	ML with region fixed effects	ML with region fixed effects	ML with region fixed effects	ML with region fixed effects
DV: Education spending per capita	Primary, secondary, tertiary	Primary, secondary, tertiary	Primary, secondary, tertiary	Primary, secondary, tertiary	Primary, secondary, tertiary	Primary, secondary, tertiary
Nested level	Region	Region	Region	Region	Region	Region
REAI Indicator	Scope*Depth	Scope*Depth	Scope*Depth	Scope*Depth	Scope*Depth	Scope*Depth
Regional cabinet ideology	24.86	2.53	5.80	16.79	15.87	-7.50
State-wide cabinet ideology	14.44	7.65	91.40***	19.54	28.06	29.15
Ideological proximity	-23.27	-11.76	-43.89**	-12.37	-8.94	-8.11
Regional cabinet ideology * Regional authority over education	**-36.68*** **	**-31.26*** **	**-31.92*** **	**-32.72*** **	**-30.09*** **	**-27.45*** **
Regional authority over education	19.56**	20.10*	16.61	17.87	17.60*	15.36
Std. of regional gov. ideology	.15	.15	.15	.15	.15	.15
Std. of state-wide gov. ideology	.13	.13	.13	.13	.13	.13
n	3483	3550	3269	3612	3517	3559
countries	14	14	14	14	14	14
Country Excluded	Germany	Italy	Japan	Norway	Spain	Sweden
regions	266	261	235	264	265	262

Controls: Duration reg. cabinet in days, Number of parties in cabinet, Percentage of population < 14, Regional population (in 100.000), Log of population density, Regional GDP per capita (in 1000), Regional GDP growth (in %), Mountain region, Multilingual education, Unemployment, Special region, Nr. of regions per country.

Table A10: *Robustness– Exclusion of Single Countries III*

	(1)	(2)	(3)
Estimator	ML with region fixed effects	ML with region fixed effects	ML with region fixed effects
DV: Education spending per capita	Primary, secondary, tertiary	Primary, secondary, tertiary	Primary, secondary, tertiary
Nested level	Region	Region	Region
REAI Indicator	Scope*Depth	Scope*Depth	Scope*Depth
Regional cabinet ideology	17.45	14.24	14.06
State-wide cabinet ideology	-2.16	17.22	18.55
Ideological proximity	26.03	-8.09	-11.58
Regional cabinet ideology * Regional authority over education	**-35.03*** **	**-33.31*** **	**-32.69*** **
Regional authority over education	21.12**	25.77	17.90*
Std. of regional gov. ideology	.15	.15	.15
Std. of state-wide gov. ideology	.13	.13	.13
n	3256	3699	3739
countries	14	14	14
Country Excluded	Switzerland	United Kingdom	Unites States
regions	259	278	282

Controls: Duration reg. cabinet in days, **number** of parties in cabinet, Percentage of population < 14, Regional population (in 100.000), log of population density, Regional GDP per capita (in 1000), Regional GDP growth (in %), Mountain region, Multilingual education, Unemployment, Special region, Nr. of regions per country.

4.5 Control-Dependency – Exclusion of Controls

After assessing the importance of single countries for our inference we turn to the importance of specific control variables. Control variables are typically very influential once they have a statistically significant effect on the main independent variable as well as the main dependent variable. Once we are interested in an interaction effect, this description has to hold true for constituent parts of the interaction. Table 1 in the paper shows the distribution of controls across left versus right leaning regional governments. Based on the rather even distribution substantial changes in the regional government ideology effects by including controls in unlikely. However, as interaction of government ideology and regional authority over education involves more variables, we assess whether the exclusion of specific controls makes a difference for our interaction of main concern. Indeed, none of the controls appears to be challenging our core findings. The biggest changes are highlighted in bold. Omitting the regional population size would have led to an underestimation of the effect whereas omitting the regional GDP would have led to an over-estimation (compare Table B7).

Table A11: *Effects of omitting single controls*

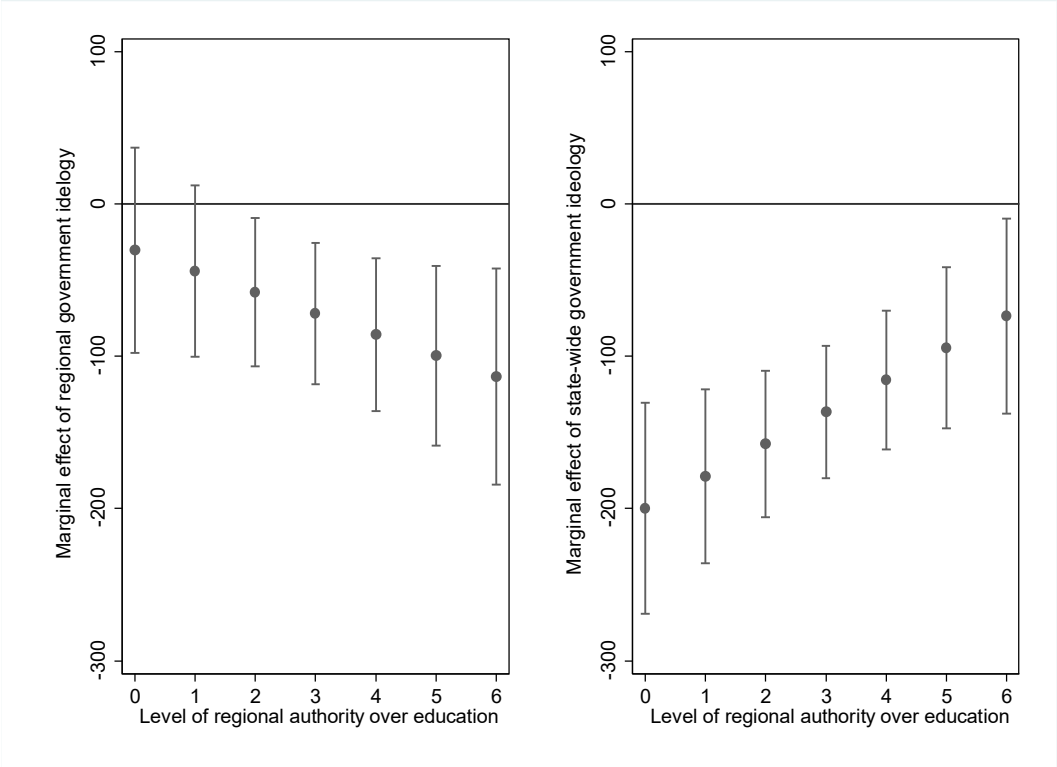
	(1)
Estimator	Panel fixed effects
Control excluded	Interaction effect (Benchmark = -32.73***)
State-wide cabinet ideology	-32.73***
Ideological proximity	-32.73***
Duration reg. cabinet in days	-32.69***
Number of parties in cabinet	-32.96***
Percentage of population < 14	-35.23***
Regional population (in 100.000)	**-29.89*** **
Log of population density	-34.83***
Regional GDP per capita (in 1000)	**-37.75*** **
Regional GDP growth (in %)	-33.12***
Mountain region	-32.69***
Multilingual education	-32.73***
Unemployment	-33.89***
Special region	-32.73***
Nr. of regions per country	-34.81***
Std. of regional gov. ideology	.15
Std. of state-wide gov. ideology	.13
n	3739
countries	15
regions	282

Note: Constituent parts of the interaction (REAI and regional government ideology) are always included.

Models with many controls can be suspicious and might in some case represent instances of over-identification. To this objection, we have two responses: The short answer is that the main findings replicate in models without any control variables (see Figure A13 below). The longer answer is that we think our manuscript actually discusses, explains, and tests already quite extensively which control variables should be included for both theoretical and empirical reasons: In Table 1 of the paper, we show the distribution of control variables between "treatment" and "control" cases, i.e. between cases with above-average and below-average ideological positions. The table indicates that differences in the results between models including controls and those without should not be too severe, as the covariates are rather balanced between both groups. Nonetheless, there are some differences. For example, regions governed by rightwing parties have a lower population density and the right tends to govern more frequently in mountainous terrains. Controlling for population density should thus decrease the regional government effect because it reduces per capita spending (economies of scale) and would, without controlling for it, wrongly attribute higher spending in more sparsely populated areas to the right. Controlling for mountains terrains should also decrease the government effect because mountainous terrain increases per capita spending for structural reasons and right leaning governments are more frequent in these areas (here structural expenses would be wrongly attributed to the right). In a model without controls we thus expect that differences between more leftist and rightist governments are reduced because structural factors which make education less costly coincide with left-wing governments but we expect left governments to raise education spending. That is, the negative effect of right-wing governments on educational expenditure would appear less negative because structural and cost-increasing factors would be "wrongly" attributed to the right. (We are also not concerned about endogeneity problems especially with these two variables, as it is unlikely that changes in regional education spending affect previous changes in population density or in the physical makeup of the landscape.)

Figure A.13 confirms this empirically by replicating our main interaction findings in a model without any control variables. The overall picture is unchanged but regional government effects are relatively weaker (for the reasons discussed above).

Figure A13: Interactions of government ideology and regional authority over education without including controls.



4.6 Alternative Ideology Scales

We deliberately selected an ideology scale which is limited to the economic dimension of political conflict. The approach of Röth (2017) has shown to be the highest validity in that respect. Alternative specifications of a general left-right dimension would include a "cultural dimension" which we perceive as hardly comparable across regions and countries for several reasons. What means "culturally left" and "culturally right" is conceptually very difficult on the state-wide level already. Adding this ambiguity to a multi-level context where traditional morals or nationalism can have many meanings once minority nationalist are part of the comparison, we think that left and right loses a core assumption for rigorous research – namely, equivalence.

Nonetheless, we check two alternative ideology scales instead of the preferred economic left right dimension. The first is the RILE (Laver and Budge 1992) scale based on the same manifesto source and including an economic as well as a cultural dimension. It is one of the most widespread applied ideological dimension of political parties. We aggregated the party position on the Rile scale in exact correspondence to the approach we applied for the economic dimension. Furthermore, we also assessed the scale of Franzmann and Kaiser (2006) which conceptually deviates from the RILE scale as it is based on the same data source such as the other two scales and argues for functional equivalence of different policy statements over time and countries. Equivalence, is thus argued to be achieved by country and time specific scale construction (see Franzmann and Kaiser 2006 for a detailed discussion and Röth 2017 for a discussion of related problems).

In Table B8 we depict the results of a comparison of different scales which is based on the exact same observations. Thus, the results differ from the models in the main text because they rest on fewer observations in order to make the alternative ideology scales comparable. The economic dimension of Röth (2017) is negative and significant in all three specifications. The Rile index show similar results but fails to reach significance in the multi-level model. The Franzmann and Kaiser (2006) approach does not show any significant results. We think, this is related to the properties of the RILE and the Franzmann and Kaiser approach in terms of comparability. Both approaches have been shown to have lower equivalence over geographical units and time than the approach of Röth (2017) and thus, induce more noise to the models.

Table A12: *Alternative ideology scales*

	(Benchmark)			**(1)**			**(2)**		
Estimator	Panel region fixed effects	Panel random effects	ML (country and region)	Panel region fixed effects	Panel random effects	ML (country and region)	Panel region fixed effects	Panel random effects	ML (country and region)
DV: Education spending per capita	Primary, secondary, tertiary	Primary, secondary, tertiary	Primary, secondary, tertiary	Primary, secondary, tertiary	Primary, secondary, tertiary	Primary, secondary, tertiary	Primary, secondary, tertiary	Primary, secondary, tertiary	Primary, secondary, tertiary
REAI Indicator	Scope*Depth	Scope*Depth	Scope*Depth	Scope*Depth	Scope*Depth	Scope*Depth	Scope*Depth	Scope*Depth	Scope*Depth
Ideology Scale	Economic (Röth 2017)	Economic (Röth 2017)	Economic (Röth 2017)	RILE (Laver and Budge 1992)	RILE (Laver and Budge 1992)	RILE (Laver and Budge 1992)	Left-Right (Franzmann & Kaiser 2006)	Left-Right (Franzmann & Kaiser 2006)	Left-Right (Franzmann & Kaiser 2006)
Regional cabinet ideology	-8.87	28.25	-18.79	142.50***	171.52***	46.61	-56.61	-96.60**	-64.72
State-wide cabinet ideology	-54.88***	6.66	-95.65***	13.46	-5.09	-104.05***	8.25	-7.95	-103.33***
Ideological proximity	56.84***	4.12	73.24***	1.71	2.53	70.29***	-11.44	-10.75	69.27***
Regional cabinet ideology * Regional authority over education	-20.65**	-42.46***	-21.04**	-46.76***	-52.48***	-9.80	-6.31	0.37	6.67
Regional authority over education	9.41	71.63***	19.51*	18.88*	72.22***	13.91	0.25	47.35	9.06
Std. of regional gov. ideology	.15	.15	.15	0.10	0.10	0.10	0.18	0.18	0.18
Std. of state-wide gov. ideology	.13	.13	.13	0.13	0.13	0.13	0.13	0.13	0.13
n	3469	3469	3469	3469	3469	3469	3469	3469	3469
countries	14	14	14	14	14	14	14	14	14
regions	257	257	257	257	257	257	257	257	257

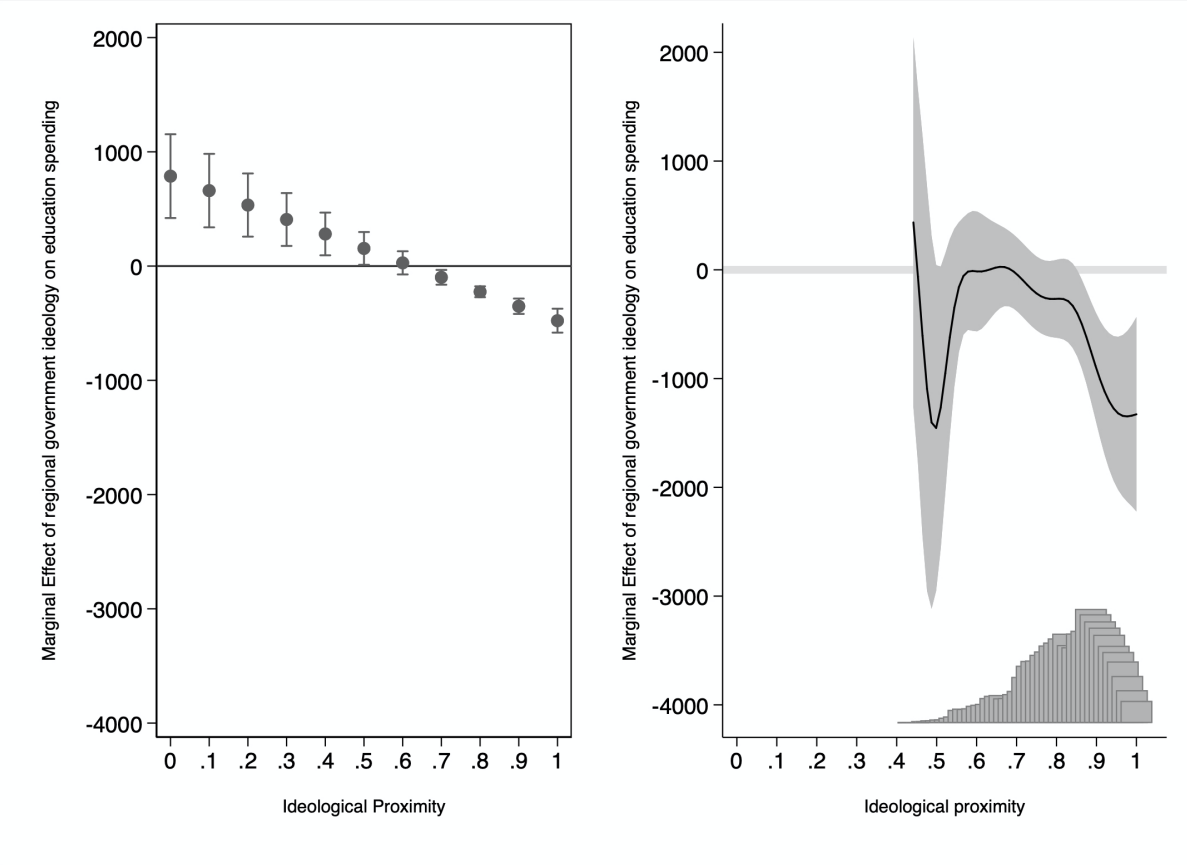
Controls: Duration reg. cabinet in days, **Number** of parties in cabinet, Percentage of population < 14, Regional population (in 100.000), Log of population density, Regional GDP per capita (in 1000), Regional GDP growth (in %), Mountain region, Multilingual education, Unemployment, Special region, Nr. of regions per country. The results are based on fewer observation than those presented in Table 1 in the main text. In order to make the effect of different ideology scales comparable we limited the sample to observations which are existent for the three scales.

4.7 Opportunity structures for ideology effects

In the final section of the robustness part we elaborate on opportunity structures which might be conducive to regional parties in government to make a difference on educational policy. Our data-set include various potential variables which could be seen as opportunity structures. In order to assess their reform facilitating or hampering effects we first limit our sample to regions where regional governments have indeed institutional leeway over education (REAI >=3) because have shown already robustly that regional governments do not matter in context of lower regional authority over education. In a second step, we interact regional government ideology on the various potential opportunity structures within the sub-sample of regions with REAI >= 3. Surprisingly, socio-economic and demographic conditions seem not to matter (Unemployment, population density, percentage of people under 14, population size, GDP growth and level, duration of government in days per year, number of parties in cabinet). None of the interactions between regional government ideology and those opportunity structures is significant (not shown). The negative impact of right-leaning regional governments on education spending is thus systematic and not systematically dependent on specific socio-demographic contexts.

The only condition of relevance is the ideological alignment between the state-wide government and the regional government. Figure B1 depicts first the linear and second the non-linear relationship between regional government ideology und different degrees of ideological proximity. The linear interaction of the left-hand side of Figure B1 is misleading, as the positive spending effect of right-leaning regional governments in context of ideological distance to the state-wide government is basically interpolated. Linear interaction effects are problematic in constellations of low overlap or in parts of the distribution where not many observations exist (Hainmueller et al. 2019). Accordingly, we use a non-parametric kernel density estimator to estimate the exact relationship between regional government ideology and education spending in different contexts of ideological proximity. The distribution of ideological proximity shows that values below 0.4 are virtually absent. As expected the positive effect in the left part of the linear interaction is based on linear interpolation but not on real observations. The non-linear interaction reveals that below a proximity value of 0.8 we do not find a significant and negative effect of right-leaning regional governments on education spending. That means, sufficient ideological similarity between the regional and the state-wide government is necessary in order to observe systematic differences on regional education spending induced by regional government.

Figure A14: *Linear and non-linear interaction of regional government ideology and ideological proximity to the state-wide government.*



Note: Linear interaction effects are problematic in constellations of low overlap or in parts of the distribution where not many observations exist (left side; see Hainmueller et al. 2019 for the argument). Accordingly, we use a non-parametric kernel density estimator to estimate the exact relationship between regional government ideology and education spending in different contexts of ideological proximity (right side).

4.8 Findings for the control variables

Besides the partisan effects, the control variables also show a range of interesting findings[2]: Interestingly, regional authority does not have a spending effect on its own. This puts doubts on the claim of previous studies hat more decentralized systems tend to exhibit higher spending levels. Furthermore, neither the coalitional size (number of parties in the government) nor government duration affects regional education spending across models.

A higher share of younger citizens, as well as more densily setteled populations decrease per capita education spending. A likely reason is economies of scale: The provision of education is easier and relatively less costly the more pupils attend education.

Mountaneous topographies substantially raise per capita costs of education. In mountains, infrastructre for education as well as lower student-teacher ratios increase per capita costs. In contrast, we do not find substantial effect of mountains on higher education spending, which makes sense as higher eduction institutions are usually situated in non-montaineous and more densily populated areas.

The level and change in regions' economic performance also affect education spending. Wealthier regions spend more on education per capita, but interestingly economic growth shows a consitent negative effect on education spending. A potential reason is that in times of economic decline more money might be spend on education to bolster unemployment. In line with this, regional unemployment are positively related to per capita spending. Thus, economic downturns seem to be less of a restriction for regional governments to curb education spending but rather an incentive to invest. Yet, for reasons of path-dependencies it might be that education spending does not entirely catch-up with growth rates and at the same time retrenchment does not comply with economic downturns.

Another final important factor is the provision of multilingual education, typically found in regions where language minorties have succesfully strived for the opportunity to be taught in their traditional mothertongues.

[2] We focus on the results of the random effects models here, as slowly changing or static variables are harder to interpret in the fixed effects models. Yet, the coefficients in both specifications are highly similar.

4.9 Why the economic dimension?

In the Appendix I part 4.6, we discussed why we used the economic dimension in more detail and assessed the results while using alternative ideology scales. We deliberately selected an ideology scale which is limited to the economic dimension of political conflict. The approach of Röth (2017; 2018) has shown to be the highest validity in that respect. Alternative specifications of a general left-right dimension would include a "cultural dimension" which we perceive as hardly comparable across regions and countries for several reasons. What means "culturally left" and "culturally right" is conceptually very difficult on the state-wide level already. Adding this ambiguity to a multi-level context where traditional morals or nationalism can have many meanings once minority nationalist is part of the comparison, we think that left and right loses a core assumption for rigorous research – namely, equivalence.

Nonetheless, we check two alternative ideology scales instead of the preferred economic left right dimension. The first is the RILE (Laver and Budge 1992) scale based on the same manifesto source and including an economic as well as a cultural dimension. It is one of the most widespread applied ideological dimensions of political parties. We aggregated the party position on the Rile scale in exact correspondence to the approach we applied for the economic dimension. Furthermore, we also assessed the scale of Franzmann and Kaiser (2006) which conceptually deviates from the RILE scale as it is based on the same data source such as the other two scales and argues for functional equivalence of different policy statements over time and countries. Equivalence, is thus argued to be achieved by country and time specific scale construction (see Franzmann and Kaiser 2006 for a detailed discussion and Röth 2017 for a discussion of related problems).

In Table B8 we depict the results of a comparison of different scales which is based on the exact same observations. Thus, the results differ from the models in the main text because they rest on fewer observations in order to make the alternative ideology scales comparable. The economic dimension of Röth (2017; 2018) is negative and significant in all three specifications. The Rile index show similar results but fails to reach significance in the multi-level model. The Franzmann and Kaiser (2006) approach does not show any significant results. We think, this is related to the properties of the RILE and the Franzmann and Kaiser approach in terms of comparability. Both approaches have been shown to have lower equivalence over geographical units and time than the approach of Röth (2017; 2018). These

properties are highly problematic in our context as we argued that institutional characteristics such as regional authority a key-concepts in our study and rarely vary within regions or countries but predominantly across countries. Since we strongly rely on cross.country variance and RILE and even to a stronger degree Franzmann/Kaiser ideology scales are hardly comparable across countries we have stronger trust in models which build on ideology scales which are comparable across countries.

One reviewer mentioned that expert placement (such as CHESS) might be a reasonable alternative. We agree in principle, but want to pint to two aspects. Expert placements are way less frequent that manifesto placements which would lead us to have either a lot of missing or a lot of imputation. For that reason, we abstain from expert placements but want to highlight that ideological measure we use has, to our knowledge, the highest convergent validity to expert placements. Table A13 depict the correlation of expert placements and different scaling approaches.

Table A13: Economic dimension versus alternative procedures and left/right

Benchmark	Economic Dimension (CHESS all waves)
Economic dimension (as used)	0.86
Economic (Franzmann and Kaiser 2006)	0.71
Left Right (Franzmann and Kaiser 2006)	0.76
Rile (Laver and Budge 1992)	0.67
n	519

Note: As benchmark we use the economic dimension of the Chapel Hill Expert Survey with the rescaled values corrected with anchoring vignettes (Bakker et al. 2014; 2015).

5. Correspondence of state-wide and regional platforms

We agree that using state-wide party manifestos to obtain regional government position is a procedure with strong assumptions. We are not about deny that regional platforms regularly diverge from the state-wide party organization. We keep using that approach for two reasons. First, there is no valid alternative yet. We simply lack comparable regional party position data on the regional level which would us allow to assess our research question. Second, differences between state-wide and their regional party branches are not as pronounced as to prohibit plausible regional government differences. Spain provides a

unique opportunity to compare state-wide and regional parties' platform based on the same data source (state-wide vis-à-vis regional manifestos) even coded by the same scheme. Accordingly, we use national party positions as based on manifesto data (Volkens et al. 2019), transformed into an economic and a cultural dimension following the procedure of Röth et al. (2018; 2016). In the cases of Spain, we use the regional manifesto data (Alonso et al. 2013). Regional manifesto data in Spain also transformed using the exact same approach.

Time points between regional and national elections differ Spain in some elections but not in others. When the year of elections on the national and regional level coincide, correlation on the economic dimension between regional and national party position of the same party is 0.88. The correlations marginally decrease if we include the closest match in terms of the election year (0.84). The differences between regional and national party platforms are more pronounced on the cultural dimension (0.65 versus 0.57 when the time of election diverges).

As discussed before, we expected comparability issues on the cultural dimension and this is confirmed by the correlations. For us, a correlation of 0.88 is sufficient to assume ideological differences between the center left and center right are revealed by our measure. Furthermore, Spain has very regionalized party system, including branches of the state-wide parties which are even organizational separate parties in the Spanish autonomous regions. That being said, we assume that correlations in other countries are likely to be at least as high as in the case of Spain.

Table A14: Correspondence of national and regional platforms in Spain

	Economic Dimension (regional)	**Cultural Dimension (regional)**
Country	**Spain**	**Spain**
Economic Dimension (national)	0.84	-
Cultural Dimension (national)	-	0.57
Temporal match	extended	extended
	N=154	**N=154**

5. References

Alonso, S., Gómez, B., & Cabeza, L. (2013). Measuring centre–periphery preferences: The regional manifestos project. *Regional & Federal Studies, 23*(2), 189-211.

Bakker, R., De Vries, C., Edwards, E., Hooghe, L., Jolly, S., Marks, G., ... & Vachudova, M. A. (2015). Measuring party positions in Europe: The Chapel Hill expert survey trend file, 1999–2010. Party Politics, 21(1), 143-152.

Bakker, R., Jolly, S., Polk, J., & Poole, K. (2014). The European common space: Extending the use of anchoring vignettes. *The Journal of Politics, 76*(4), 1089-1101.

Franzmann, S., & Kaiser, A. (2006). Locating political parties in policy space: A reanalysis of party manifesto data. *Party politics*, *12*(2), 163-188.

Hainmueller, J. (2012). Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies. *Political Analysis, 20(*1), 25-46.

Hainmueller, J., & Xu, Y. (2013). Ebalance: A Stata package for entropy balancing. *Journal of Statistical Software, 54*(7).

Hooghe, L., Marks, G., Schakel, A. H., Osterkatz, S. C., Niedzwiecki, S., & Shair-Rosenfield, S. (2016). *Measuring regional authority: A postfunctionalist theory of governance* (Vol. 1). Oxford University Press.

Laver, M. J., & Budge, I. (1992). Measuring policy distances and modelling coalition formation. In *Party policy and government coalitions* (pp. 15-40). Palgrave Macmillan, London.

Röth, L. (2017, June). Equivalence presupposes validity. Toward comparable party positions on the market dimension. In EPSA Conference (pp. 22-24).

Röth, L., Kaiser, A., Varol, Ç., & Sadioğlu, U. (2016). Centre-periphery conflict and ideological distance in Turkey. *Swiss Political Science Review, 22*(4), 585-607.

Röth, L., Afonso, A., & Spies, D. C. (2018). The impact of Populist Radical Right Parties on socio-economic policies. *European Political Science Review, 10*(3), 325-350.

Röth, L. (2018). Political parties and the market. Towards a Comparable Assessment of Market Liberalism. University of Cologne.

Volkens, A., Burst, T., Krause, W., Lehmann, P., Matthieß T., Merz, N., Regel, S., Weßels, B., Zehnter, L. (2020): The Manifesto Data Collection. Manifesto Project (MRG/CMP/MARPOR). Version 2020a. Berlin: Wissenschaftszentrum Berlin für Sozialforschung (WZB). https://doi.org/10.25522/manifesto.mpds.2020a