Waving goodbye? The determinants of autonomism and secessionism in Western Europe

Brais Álvarez Pereira, Martín Portos and John Vourdas

ABSTRACT

Waving goodbye? The determinants of autonomism and secessionism in Western Europe. *Regional Studies*. This paper sheds light on the main aggregate-level determinants of electoral support for regionalist parties across 10 Western European countries. A region being relatively richer than the country to which it belongs is associated with higher electoral support for regionalist parties only to the extent that the region is culturally differentiated. This hypothesis is substantiated theoretically, tested empirically and found to hold in the form of a strong and significant interaction effect between cultural and economic variables. This result, omitted in previous studies, implies a profound change in the interpretation of the role of income and cultural differences in explaining support for regionalism, for both autonomist and separatist parties.

KEYWORDS

secessionism; autonomism; regionalism; Western Europe; cultural distance; electoral behaviour

摘要

挥手告别？西欧自治主义与分离主义的决定因素。*Regional Studies*. 本文对于西欧十个国家选举支持区域政党的总体层级之主要决定因素提出洞见。一个较其所在国家更为富裕的区域，在选举上更加支持区域主义政党，但仅在该区域在文化上具有差异时此种假设具有理论根据，且进行了经验鉴定，并发现支持文化与经济变项之间强烈且显著的互动效应形式。在过往研究中受到忽略的此般结果，意味着所得与文化差异，在解释对区域主义与主张自治和分离主义政党的支持中所扮演的角色有着深刻的改变。

关键词

分离主义; 自治主义; 区域主义; 西欧; 文化距离; 选举行为

RÉSUMÉ

Dire au-revoir? Les facteurs déterminants de l'autonomisme et du sécessionnisme en Europe occidentale. *Regional Studies*. Cette communication présente les principales causes du soutien électoral au profit des partis régionalistes de 10 pays d'Europe occidentale. Le fait qu'une région soit relativement plus riche que le pays auquel elle appartient est lié à une participation électorale importante en faveur des partis régionalistes, seulement dans la mesure où la région se distingue sur un plan culturel. Cette hypothèse est étayée théoriquement, testée de façon empirique, et s'avère fiable sous forme d'un effet interactif significatif entre des variables culturelles et économiques. Ce résultat, qui avait été omis de précédentes études, implique une profonde évolution dans l'interprétation du rôle des revenus et des différences culturelles, pour expliquer le soutien pour le régionalisme, pour des partis tant autonomistes que séparatistes.

MOTS-CLÉS

sécessionnisme; autonomisme; régionalisme; Europe occidentale; distance culturelle; comportement électoral

ZUSAMMENFASSUNG

Ein Lebewohl? Die Determinanten von Autonomismus und Sezessionismus in Westeuropa. *Regional Studies*. In diesem Beitrag untersuchen wir auf zusammengefasster Ebene die wichtigsten Determinanten der Wählerunterstützung für regionalistische Parteien in zehn westeuropäischen Ländern. Wenn eine Region relativ reicher ist als das Land, zu dem sie gehört, fällt die Wählerunterstützung für regionalistische Parteien nur in dem Maße höher aus, in dem die Region kulturell differenziert ist. Diese Hypothese wird auf theoretische Weise belegt sowie empirisch überprüft und bestätigt, wobei festgestellt wird, dass eine starke und signifikante Wechselwirkung zwischen den kulturellen und ökonomischen Variablen vorliegt. Dieses in früheren Studien übersehene Ergebnis führt zu einer grundlegenden Veränderung der Interpretation der Rolle von Einkommen und Kulturunterschieden zur Erklärung der Unterstützung von Regionalismus sowohl in Bezug auf autonomistische als auch auf separatistische Parteien.

SCHLÜSSELWÖRTER

Sezessionismus; Autonomismus; Regionalismus; Westeuropa; kulturelle Distanz; Wählerverhalten

RESUMEN

¿Un adiós? Los determinantes del autonomismo y secesionismo en Europa occidental. *Regional Studies*. En este trabajo analizamos los principales determinantes de nivel global del apoyo electoral a partidos regionalistas en diez países de Europa occidental. Si una región es relativamente más rica que el país al que pertenece, se relaciona un mayor apoyo electoral a partidos regionalistas solo si existe una diferenciación cultural. Esta hipótesis sustentada teóricamente se comprueba empíricamente, y se verifica a través de un efecto interactivo fuerte y significativo entre las variables económicas y culturales. Este resultado, omitido en estudios previos, implica un profundo cambio en la interpretación del rol de la riqueza y las diferencias culturales para explicar el apoyo a partidos regionalistas, tanto autonomistas como secesionistas.

PALABRAS CLAVES

secesionismo; autonomismo; regionalismo; Europa occidental; distancia cultural; comportamiento electoral

INTRODUCTION

The study of the determinants of the break-up of countries and the origin of new nation-states received considerable attention following the collapse of the Soviet Union and Yugoslavia.[1] Recently, this topic has regained momentum as two large regions have been pushing for their independence.[2] In Scotland the Scottish National Party (SNP)-led government held a referendum on independence from the UK in September 2014. Despite the unionist win, a sizable proportion of 44.7% of voters opted for secession (turnout of 84.59%). Likewise, Catalan authorities strove to hold an official plebiscite on independence, but faced outright opposition from the Spanish government and the constitutional court. Given these circumstances, a large-scale *non-binding* unofficial voting performance led by civil society organizations took place instead, with the collaboration of the Catalan regional government. In total, over 2.3 million votes were cast in November 2014 (the estimated turnout was approximately 40%), with 80.8% of participants ticking the 'yes–yes' option (i.e., yes for a state for Catalonia; and yes for an independent state) (Della Porta, O'Connor, Portos, & Subirats, 2017). While these are arguably two extreme cases, support for greater regional[3] autonomy has been rising within European Union member states, against a backdrop of the wider European integration process between member states. However, electoral research tends to overlook the relevance of the centre–periphery axis on vote choice (Medeiros, Gauvin, & Chhim, 2015).

In line with the most recent literature, we will refer to *regionalism* as the process of agitation within a sub-state territorial unit or region that leads to increasing demands for greater autonomy, which seeks to recognize, empower and reinforce its structure of self-government (de Winter, 1998; Massetti, 2009). Within regionalist parties, we define a party as *autonomist* if it does not intend to secede in any way from the current state, it respects the unity of the current state, and simply has an ambition for a larger degree of regional self-government (Jolly, 2006; Massetti & Schakel, 2013, pp. 5, 13). A party is *secessionist* if it promotes a territorial sub-unit's detachment from the nation-state of which it is currently part, and advocates for the right to self-determination, either by becoming an independent state or by building a new one together with other territorial units (Sorens, 2005, 2008, 2012). Consequently, we include irredentist parties as a subcategory of secessionist parties. Hence, regionalism embeds both autonomism and secessionism, with the latter being a more radical form of regionalism and the former a milder form.

Despite the rise of regionalist parties and movements in the last decades (in France, Spain, Italy and the UK, amongst others), there have been surprisingly few empirical studies on the determinants of their support. Even though there are some notable exceptions to this trend (e.g.,

Sorens, 2005, 2008, 2012; Massetti & Schakel, 2013; Brancati, 2014), we consider that two crucial issues warrant further study, as there is no consensus about them in existing literature. Firstly, what is the relation between cultural and economic factors in explaining the electoral success of different regionalist forces? Secondly, what is the connection between the two main forms of regionalism, i.e., autonomism and secessionism? Are they two different versions of the same phenomenon or two distinct phenomena?

We argue that cultural proximity accounts for the support for increasing regional autonomy through an interaction between cultural and income-related variables which has not been explored in the existing literature. The main finding is that a region being richer relative to the rest of the country is only correlated with support for more regional autonomy if its culture is significantly different from the main one in the country, and not otherwise. Similarly, the cultural–identity gap is significantly more important in accounting for the success of regionalist and secessionist parties in relatively rich regions, playing a much smaller role in poorer ones. This common feature underpins both autonomist and secessionist electoral support. Finally, the paper develops the idea that – at least in electoral studies – these two forms of regionalism are intimately intertwined, both on theoretical and empirical grounds.

The paper is structured as follows. The next section reviews the relevant contributions in the literature. The third section develops our theoretical framework. The data, empirical model and main results from the empirical enquiry are described in the fourth section. The fifth section concludes by highlighting the main implications of this paper.

LITERATURE REVIEW

The main factors identified in the literature as accounting for regionalist support – either in its autonomist or in its secessionist versions – can be divided in two broad categories: cultural and economic (Brancati, 2014; Hooghe, Marks, & Schakel, 2010; Jolly, 2006; Massetti & Schakel, 2013; Sorens, 2005, 2008, 2012).

Cultural explanations of support for greater regional autonomy emphasize the role of historical, traditional, linguistic and ethnic factors. These are often associated with some distinctive identity features, such as having a specific minority language (Fearon & van Houten, 2002; Van Houten, 2000).[4] For example, cultural explanations help to explain the support for secessionism in Catalonia and the Basque Country. These two regions have a strong tradition of identity nationalism linked to their language, and in both regions secessionist parties gathered approximately 60% of the votes in the last regional elections.[5] However, Catalan is also spoken in the regions of Valencia and the Balearic Islands, which are both autonomous communities that have traditionally shown active hostility to secessionism, although there has been some, rather mild, support for autonomism (Riera, 2014).[6] Similarly, while Scottish separatism has been on the rise in the last decades, Scottish

Gaelic is barely spoken by 1% of Scots. This contrasts with the situation in Wales, where around 20% of the population speaks Welsh, but support for secessionism is much lower. Having a distinctive minority language might not be the only (or even the main) sign of cultural difference. Other cultural or identity aspects might play a role in determining support for greater regional autonomy. In fact, Desmet, Le Breton, Ortuño-Ortín, and Weber (2011) show that physical distance between the capitals of two countries is a better proxy for cultural difference than the linguistic distance between their main languages. Yet, cultural factors alone cannot explain support for greater regional autonomy. We argue that cases of culturally differentiated, but relatively poor, regions, such as Galicia in Spain, Sicily in Italy or the French overseas departments, where the support for regionalism is relatively low or non-existent, show that cultural-identity factors alone cannot explain a large support for autonomist or secessionist parties.

It is well known that economic variables such as income, wealth and relative economic well-being may also influence secessionist and autonomist support. Appeals to primordial past, history and traditions of a nation seem to affect voting to a much lesser extent than material concerns and socio-economic issues, at least when it comes to endorsing secessionist parties (Sorens, 2005). Accordingly, debates on territorial separation (and also current pro-secessionist referendum campaigns, such as the recent Scottish and Catalan ones) have often focused on the economic efficiency losses and gains of separation vis-à-vis the status quo.[7] From an individual level of analysis, political economists have addressed this issue as well. The established literature on the break-up of nation-states, for instance, offers a trade-off that determines whether or not secession is optimal for a given individual (e.g., Alesina & Spolaore, 1997). While integration facilitates coordination and minimizes the duplication of fixed costs in the provision of public goods (e.g., in defence and law enforcement), the main incentive for independence is that the benefits from remaining united are generally not evenly distributed among all citizens (nor regions) at the country level. For instance, heterogeneity in tax contributions across regions, resulting from differences in average income per capita, accentuate economic grievances (Bolton & Roland, 1996, 1997). The latter are at the core of claims for greater autonomy and secession: the argument is that the centre's economic policies are holding the growth of the region down or that the (richer) region is subsidizing poorer ones (Horowitz, 1981). The Italian Lega Nord political party often puts these arguments forward, for example. There seems to be a widespread agreement in the literature that richer regions tend to be more prone to back autonomist (de Winter, 1998; Gordin, 2001; Van Houten, 2000) and secessionist parties (Sorens, 2005).

Some existing contributions have already considered the interrelated effect of both cultural and economic variables to explain secessionist support. For example, in their study on the trade-offs between the benefits of large jurisdictions and the costs of heterogeneity of large and

diverse populations, Alesina and Spolaore (1997) present a model that implies that democratization leads to secessions, there is an inefficiently high number of countries in equilibrium, and the equilibrium number of countries increases in the amount of economic integration. In his study of ethnic conflict in Eurasia, Hale (2008) finds an association between ethnic distinctiveness and separatism that he contends is because ethnicity accentuates difficulties of union governments at portraying themselves as non-exploiters of minority groups on the benefit of the dominant group. Focusing on Yugoslavia, Desmet et al. (2011) find that agents' preferences over different geographical configurations (i.e., the likelihood of secessions and unions) is determined by a trade-off between the costs of greater cultural heterogeneity and increasing returns in the provision of public goods.

Whilst these contributions are enlightening and advance theories to an, arguably limited, extent consistent with our main argument (i.e., economy and culture interact to promote greater support for autonomism and secessionism), this is, to the best of our knowledge, the first empirical paper with aggregated panel data that focuses on unravelling the crucial interaction effect between cultural and economic variables in explaining the support for autonomist and secessionist parties. This points not only towards an inadequacy of additive models to test for the determinants of regionalist support, but also a potential lack of theorization on the interplay between economic and cultural arguments as explanatory factors.

Another important question refers to the connection between autonomism and secessionism. It is widely accepted that these two processes are interrelated in conceptual and empirical terms (e.g., Jolly, 2006). They share many crucial aspects, as both reject and react against the status quo in favour of some alternative involving more self-government. However, some studies have stressed their diverging features (Sorens, 2008). From this point of view, autonomism and separatism would be different processes, driven by different dynamics and factors. Existing empirical evidence suggests that cultural variables are most important in determining autonomist support (Fearon & van Houten, 2002; Gordin, 2001; Hearl, Budge, & Pearson 1996), whereas economic variables are most important in determining secessionist support (Brancati, 2014; Sorens, 2005). We argue that both cultural and economic factors and, most importantly, their interaction drive the electoral support for both autonomist and secessionist parties in the same direction. This, together with other empirical findings, suggests that, at least in electoral studies, autonomism and secessionism should be considered as two varieties of the same process, namely regionalism.

THEORETICAL FRAMEWORK

Cultural proximity, relative income and electoral support for regionalism

Cultural proximity influences the feeling of belonging to a large social group and determines whom people in a given

region care about. This has an important role in the political participation process. When choosing whether or not to support greater regional autonomy, people whose region's culture is distinct from the majority of the nation will care more about the implications of that decision for the population in that same territorial subunit. On the other extreme, in regions whose culture mainly coincides with the main one in the country, people will tend to care relatively more about the effects that a certain political choice would have for the overall population in the country as a whole. In this context, this gap between the regional and the majoritarian culture in the country is known as *cultural distance*.

This effect of culture on electoral support for regional autonomy has important implications, as it introduces an important interaction between economic and cultural factors. With people in more differentiated regions caring mainly about the well-being of those in their same region, support for regionalism among the population will be large if this is thought to improve the situation of the full group, provided that the region is rich. Alternatively, the population in a relatively rich region will support regionalism to a large degree only when the region's culture is significantly different from the predominant one in the country. Those in highly differentiated but poor regions will tend to see the less decentralized status quo as better for their region's material interests, and we would expect a relatively low proportion of the electorate to support increasing regionalism. This analysis brings us to the main hypothesis of the paper:

> *Hypothesis 1: There exists a positive interaction effect between cultural and economic factors on the support for increasing regional autonomy.*

This interaction between economic and cultural aspects underpinning the vote for regionalist parties has been overlooked in the literature to date, as existing aggregate studies are based on additive models. This interaction has important implications, both to explain the support for regionalist parties and to contribute to a better understanding of the determinants of the relationship between the electoral support for secessionist and for autonomist options.

Cultural proximity, autonomism and secessionism

The interrelationship between secessionism and autonomism warrants further analysis. From a theoretical point of view, we approach this puzzle using a simple model that builds on Sorens's (2004) contribution. Voters are located on a continuum ranging from 0 to 6, with 0 representing maximum unionism and 6 representing independence. We assume that the election is regional (and not a referendum), and we care about any voter i, not necessarily the median voter. Let us assume, for simplicity, that there are only two parties in the election, a centralist one and a regionalist one, with proposed policies respectively being P_c and P_r. Voter i preferences

are given by the negative loss function:

$$U(P_j, I_i) = -a(|P_j - I_i|)^2,$$

where a is a positive constant; $U(P_j, I_i)$ is the utility of voter i under policy P_j; and I_i is the ideal policy point of voter i. In our example, the policy offered by the centralist party is $P_c = 1$, and voter i preferred policy point $I_i = 4$. If the regionalist party is either an autonomist party with $P_r = 3$, or a secessionist party with $P_r = 6$, voter i will choose to vote for the regionalist party (one of the two existing in this region). This is consistent with the fact that not necessarily all the people who vote for secessionist parties in a regional election want their region to break away from their current nation-state. Similarly, in regions without secessionist parties, a part of the electorate of autonomists parties might want the region to secede. This misalignment between the preferred level of regional autonomy and voting behaviour derives from the existence of a discrete number of parties, and raises the following question: should the determinants of autonomism and secessionism through electoral results be studied as two different processes or as two aspects of the same phenomena?

We argue that analyzing electoral support for these two options as different processes might lead to misleading conclusions. If secessionist parties are included in the group of 'non-autonomist parties' when empirically analyzing the determinants of electoral support for autonomist parties, the results might be biased (see Figure A1 in Appendix A in the supplemental data online). As a given predictor might have an effect not only on autonomism but also on the support for secessionist parties, considering only autonomist parties as an independent category biases the estimated effect of this variable downwards. Alternatively, if support for autonomist parties is put together with non-regionalists when studying support for secessionist parties, the estimate will be biased.

In order to test this, we split the full sample into two separate ones. The first of these keeps only those regions and periods that have only votes for secessionist and non-regionalist parties, or for non-regionalist parties alone. The second contains those observations in which there is a positive proportion of votes either for autonomist and non-regionalist parties, or for just non-regionalist parties. The observations for those regions that have a positive amount of votes for the two different categories of regionalist parties are excluded in both samples, i.e., regions with support for autonomist parties are excluded from the first sample and observations with support for secessionist parties are dropped in the second one. Given that we argue the same mechanism underlies electoral support for autonomist and secessionist parties, we would expect the following hypotheses to hold:

Hypothesis 2a: For the models explaining electoral support for either autonomist or secessionist parties, the coefficients for the same variables across models (with full and the separated samples) should show statistically different results.

Hypothesis 2b: When only the separated samples are considered, coefficients explaining support for autonomism and secessionism should not differ in statistical terms.

If hypotheses 2a/2b are satisfied, placing secessionism and autonomism in the same continuum measuring different degrees of regionalism will be justified, with maximal unionism and secessionism located at the extremes, and autonomism in between. As both increasing regional autonomy and secessionism imply steering policies to be more regional specific, we would expect the effect of the interaction between economic and cultural variables to go in the same direction – albeit not necessarily with the same intensity – for the two forms of regionalism. This gives the third hypothesis of the paper:

Hypothesis 3: The interaction effect between cultural and economic variables is positive in models explaining the electoral support for both autonomism and secessionism separately.

Confirming this hypothesis (with the separated samples) is consistent with the view of secessionism and autonomism as two forms of minority nationalism.

EMPIRICS

Data

We built a cross-sectional dataset using data from regions of Portugal, Spain, France, the UK, Italy, Germany, Belgium, the Netherlands, Denmark and Austria (NUTS-1 regions for Germany, Belgium, Denmark and the UK; NUTS-2 for the other cases; see Table A1 in Appendix A in the supplemental data online).

We then merged it with the 1969–2000 dataset from Sorens (2005), updating this with the most recent elections – till 2014 – and adding some explanatory variables and regions.[8] Our analysis was restricted to West European countries for two reasons. Firstly, they all have established and comparable regional elections. Secondly, while cultural differences are highly correlated with geographical distances across Western Europe (Desmet et al., 2011), this is arguably not the case in countries that were settlement colonies in the past. By reducing the sample to Western Europe, we can confidently use geographical distance as a proxy for cultural differences.

The dependent variables are respectively given by the percentage of votes for autonomist (*v_auton*), secessionist (*v_sec*), and the sum of both types of parties (*v_reg*) in a given regional election.[9] As the centre–periphery territorial cleavage tends to be channelled through the institutional domain in well-established democracies, these are used as proxies for general support for each of these political options.

To determine under which category a given party falls, we use different secondary and primary sources (manifestos, party websites and various academic contributions). Our party classification fits Nordsieck's (2014) recent categorization for that respective election (see Table A2 in Appendix A in the supplemental data online).[10] For

instance, the autonomist Partido Regionalista de Cantabria is in favour of greater regional autonomy for Cantabria but not independence from Spain, whereas the Scottish SNP is characterized as secessionist because it campaigns on a platform seeking independence from the UK.[11]

Our two main groups of explanatory factors, cultural proximity and economic aspects, are operationalized through different predictors, as follows.

Cultural variables

- *Language* is the proportion of residents in the region who speak a language other than the predominant one in the country. Having a regional vernacular language strengthens the regional cultural identity.[12] Aiming at protecting it, citizens may be keener to vote for regionalist parties (data sources: Ethnologue, 2013; Sorens, 2005).[13]
- *Distance* is the straight-line physical distance between the capitals of the region and the country, measured in hundreds of kilometres. In Western Europe, physical distance has historically been an obstacle to ensuring political coordination, cultural homogenization and reinforcing a sense of belonging. As Desmet et al. (2011) show using World Values Survey (WVS) data, the correlation between cultural distance and geographical distance is high (0.52) between West European countries.[14] As a proxy for cultural distance, we expect *distance* to be positively correlated with a region's desire for greater regionalism – and the effect should be larger for relatively rich regions.

A concern is whether *distance* might affect *v_reg* significantly through alternative channels, other than cultural. As abovementioned, using geographical distance as a proxy for cultural distance would be problematic for the whole world, particularly for former settlement colonies (e.g., United States, Canada, Australia). We test for the effect of *distance* on *v_reg* in these three large former colonies. The coefficient for this variable and the interaction *distance_GDP* (where GDP refers to gross domestic product ratio) are not significant in any of the regressions, pointing towards alternative channels through which geographical distance could affect electoral support for regionalist parties not being particularly relevant.[15] As an alternative test, we run a two-stage least squares (2SLS) model to explain *v_reg*, instrumenting *distance* and its interaction *distance_GDP* with language and its interaction *language_GDP*. Both variables are positive and statistically significant (see Table A6 in Appendix A in the supplemental data online).[16]

Economic variables

- *GDP_ratio* measures the relative difference in the average income between the region and the whole country (regional GDPpc/national GDPpc). We centre this variable around 1, such that it takes the value 0 when the GDP per capita of the region and that of the

country are equal, and positive (negative) if the GDP per capita of the region is larger (smaller) than that of the nation as a whole.[17] We expect a positive effect for this variable through its interaction with cultural distance (with positive coefficients for the interactions *distance_GDP* and *language_GDP*): richer regions will support greater autonomy and secession to avoid redistribution to poorer regions when cultural distance is larger.
- *Richer_neighbour* is a dummy that captures whether the region has a border with another country that is richer than the country to which it belongs, as this is associated with positive economic expectations toward independence or greater autonomy.

We also include different control variables in our models, as follows.

Main controls

- As support for greater regional autonomy might have increased over time, *year* captures the year when the election took place.
- *Has_neighbour* is a dummy (taking the value 1 if the region has a border with another country, and 0 otherwise) to isolate the effect of *richer_neighbour* from the general effect of having a neighbour.
- *Population*: The number of the region's inhabitants, in hundreds of thousands. The impact of having a larger population on the desire for greater autonomy and secession is ambiguous. More populated regions tend to have more weight on countrywide political decisions, which should decrease the support for increasing regional autonomy. However, large populated regions are more viable as independent states, which should increase the support for secessionism.
- *Chamber_size* measures the number of seats in the regional parliament, and is positively associated with multipartism.
- *Regional_elect* is a dummy that distinguishes the nature of the election in Sorens's (2005) data (1 = regional; 0 = national election). We expect voters to support regional parties more strongly in regional elections.
- *Country FEs* (fixed effects) are country dummies that capture specific (institutional, historical, cultural) characteristics of each country that might be correlated with support for regional parties.

Besides these variables (see Table 1 for descriptive statistics; for the correlation between the main predictors, see Table A3 in Appendix A in the supplemental data online), additional controls are used for robustness checks. We control for the level of regional autonomy (*regional_auton*), irredentist potential (*irredentist_potential*), whether the region is part of a European Union member state (*EU*), it has past historical records of independence (*hist_together*; *indep*) and is the capital of the country (*Capital*), religious differences (*religion*), squared physical distance (*distance_sq*), the degree of multipartism

Table 1. Descriptive statistics.

Variable	N	Mean	SD	Minimum	Maximum
GDP_ratio	1232	−0.041	0.222	−0.597	1.197
distance	1309	3.240	5.159	0	93.755
language	1309	0.142	0.241	0	1
distance_GDP	1232	−0.279	2.263	−40.144	15.739
language_GDP	1232	−0.009	0.082	−0.408	0.443
year	1309	1,990.937	9.243	1969	2014
has_neighbour	1309	0.503	0.500	0	1
neighbour_richer	1309	0.254	0.435	0	1
population	1269	25.555	39.533	0.298	492.71
regional_elect	1309	0.415	0.493	0	1
chamber_size	1179	67.256	39.489	2	237
ENERP[a]	1073	3.673	1.400	0.529	10.721
hist_together[a]	1173	0.559	0.490	0	1
region_auton[a]	1309	2.183	1.239	0	4
irredentist_potential[a]	1309	0.112	0.315	0	1
party_elsewhere[a]	1309	0.133	0.339	0	1

Note: [a]Variables only used in robustness checks.

(*ENERP*), and if some of its regional parties also contest elections in other regions (*party_elsewhere*) – for a description of these variables, see Appendix B in the supplemental data online.

Although all these variables have been tested, some are excluded from the reported models for various reasons, such as having important endogeneity problems (e.g., the degree of autonomy already obtained by the region and *ENERP*) or not being significant in several of the specifications. The latter includes some variables described above, such as *distance_sq*, *EU*, *religion*, *indep* and *capital*, as well as other minor controls such as the legal status of the language (whether or not it is official), whether or not the region is in the mainland, population density and ideological gap, measured as the absolute difference in votes for left-wing parties in the territorial subunits for their respective last regional and general elections.

The empirical model

The estimation methodology used is a pooled ordinary least squares (OLS) regression with country fixed effects and errors clustered at the regional level.[18] This equation gives the main model specification:

$$Votes_{jkt} = \propto_k + \beta GDP\_ratio_{jt} + \delta distance_j + \rho language_j$$
$$+ \lambda (distance * GDP\_ratio)_{jt}$$
$$+ \gamma (GDP\_ratio * language)_{jt} + \sum_{j=1}^{H} \theta_h X_{jt} + \in_{jt}$$

The percentage of votes for regionalist parties in region *j*, country *k* and period *t* is modelled as a function of *GDP_ratio*, *distance*, *language*, the interactions *distance_GDP*, *language_GDP*, and other controls X. As some model specifications include country dummies, there is

a potential dependence of constant $\propto_k$ on the country *k*, i.e.,

$$\propto_k = \propto + \sum_{k=1}^{K-1} \omega_k 1_{1\epsilon k}$$

where $1_{1\epsilon k} = 1$ if region *i* belongs to country *k*, and 0 otherwise.

Although *v_reg* is the main dependent variable, electoral support for the two subtypes of regionalist parties, autonomists (*v_auton*) and secessionists (*v_sec*), are also used as dependent variables in some model specifications.

The main hypotheses from the theoretical framework are tested empirically in the following manner:

- Hypothesis 1: we expect that (at least one of) the coefficients for *distance_GDP* and *language_GDP* will be positive and significant. Formally $H_0 = \lambda > 0, \; \gamma > 0$.
- Hypothesis 2a: a Wald test comparing the coefficients of the models for *v_sec* using the full and separated samples should report statistical differences. The same holds for *v_auton*.
- Hypothesis 2b: in the models for *v_auton* and *v_sec* using separated samples, the main coefficients should not be statistically different.
- Hypothesis 3: hypothesis 1 should hold also when explaining support for autonomist and secessionist parties separately. As we argue using the full sample would yield biased coefficients, we use the separated samples. At least either λ or γ should be significant in the corresponding regressions.

Results

Table 2, with the pooled OLS analysis of the determinants of *v_reg* (large sample, errors clustered by region), gives the main results of the paper. The richest specification,

Table 2. Vote for regionalist parties.

	(1) v_reg	(2) v_reg	(3) v_reg	(4) v_reg
GDP_ratio	19.82***	0.291	13.90***	−3.288
	(6.044)	(5.181)	(4.093)	(4.630)
distance	0.767***	1.695***	0.683***	1.528***
	(0.252)	(0.320)	(0.215)	(0.305)
language	23.09***	28.16***	17.03***	20.44***
	(7.464)	(5.950)	(6.181)	(5.915)
distance_GDP		2.636***		2.184***
		(0.736)		(0.676)
language_GDP		52.71***		43.66***
		(17.46)		(15.51)
year			0.152***	0.158***
			(0.0412)	(0.0387)
has_neighbour			−2.763	−4.894*
			(2.932)	(2.691)
richer_neighbour			12.63***	11.54***
			(3.339)	(3.235)
population			−0.162***	−0.165***
			(0.0582)	(0.0511)
chamber_size			0.0849*	0.113***
			(0.0493)	(0.0419)
regional_elect			2.681***	2.453***
			(0.677)	(0.697)
Country fixed effects	No	No	Yes	Yes
Observations	1232	1232	1102	1102
R^2	0.317	0.420	0.531	0.587
Adjusted R^2	0.316	0.417	0.523	0.579

Notes: Pooled ordinary least squares (OLS) model explaining votes for regionalist (autonomist + seccesionist) parties using the full sample. The constant term is excluded for paucity reasons. Errors are clustered at the regional level. Standard errors are given in parentheses.
*$p < 0.10$, **$p < 0.05$, ***$p < 0.01$.

including country fixed effects, is our preferred model (column 4). Overall, the solid empirical evidence provided – supported by numerous robustness checks – confirms our four hypotheses. We analyze our results for each hypothesis next.

Hypothesis 1

The interaction terms *distance_GDP* and *language_GDP* have a positive effect on *v_reg* at the 1% level (Table 2). When the interactions are introduced into the model, *GDP_ratio* is no longer significant. This means that relative regional per capita income accounts for regional party support through its interaction with cultural variables, and not by itself. Without the interaction terms, this coefficient is significant (columns 1 and 3, Table 2) because of omitted variables bias. This interaction, which has been overlooked to date in empirical studies, is crucial for interpreting the main factors behind the support for regionalist parties.

To illustrate this, we analyze the marginal effects of this interaction in our preferred model (column 4, Table 2). If a

region is 100 km farther away from the capital of the country than another region, *ceteris paribus*, the expected difference in the percentage of votes cast for regionalist parties depends on *GDP_ratio* as given by:

$$\frac{\mathrm{d}\, v\_reg}{\mathrm{d}\, distance} = 1.53 + 2.18\, GDP\_ratio$$

Similarly, conditional upon the value of *GDP_ratio*, the effect of a marginal change in the proportion of speakers of a regional distinctive language on *v_reg* is given by:

$$\frac{\mathrm{d}\, v\_reg}{\mathrm{d}\, language} = 20.44 + 43.66\, GDP\_ratio$$

Table 3(a) shows the marginal effects of distance for different values of *GDP_ratio*. From the first row, we can see that a region located 600 km away from the capital is, *ceteris paribus*, more likely to show a greater support for regional parties than one which is 500 km away. However, this effect is conditional on the relative economic affluence of the region. If the two hypothetical regions we are

Table 3. (a) Marginal effect of increasing distance by 100 km and language from 0 to 1 on support for regionalist parties for different values of *GDP_ratio*; and (b) marginal effect of *GDP_ratio* on support for regionalist parties for different values of distance and language.

(a)

GDP_ratio	−0.5	0	0.5	0.8
$\frac{\partial v\_reg}{\partial distance}$	0.44	1.53	2.62	3.28
$\frac{\partial v\_reg}{\partial language}$	−1.39	20.44	42.27	55.37

(b)

$\frac{\partial v\_reg}{\partial GDP\_ratio}$	Language			
	0	0.25	0.5	1
0 km	0	10.91	21.83	43.66
100 km	2.18	13.09	24.01	45.84
500 km	10.92	21.83	32.75	54.58
1000 km	21.84	32.75	43.67	65.50
2000 km	43.68	54.59	65.51	87.34

comparing are relatively poor, with the GDP per capita in the region being only half that of the country (*GDP_ratio* = −0.5), being 100 km farther away implies an increase of 0.44% in votes casted for regionalist parties. If these two regions were very rich, with the GDP per capita of both regions being 80% larger than the national GDP per capita (*GDP_ratio* = 0.8), being 100 km farther away would imply a 3.28% increase in regionalist support. Similarly, if a region's per capita GDP equals the national average (*GDP_ratio* = 0), there will be a 20.44% increase in votes for regionalist parties if everybody speaks a vernacular language relative to not having a regional language. If these two same regions are very rich in relative terms (*GDP_ratio* = 0.8), regionalist electoral support will increase by 55.37%. However, if these regions are very poor (*GDP_ratio* = −0.5), the marginal effect of language on *v_reg* becomes negative. Hence, the model predicts that very poor and culturally differentiated regions will tend to be more centralist than equally poor regions which are more culturally similar to the rest of the country.[19]

The implications of the interaction term for the marginal effect of differences in GDP per capita between the region and the country are even more critical. Since the coefficient for *GDP_ratio* is not significant for any of our models with interactions, the marginal effect of a one-unit change in *GDP_ratio* on *v_reg* is fully dependent on the value taken by the cultural variables, being 0 for a region whose culture is identified as the main culture in the country.[20] Formally:

$$\frac{d\,v\_reg}{d\,GDP\_ratio} = 2.18\;distance + 43.66\;language$$

Table 3(b) shows how the effect of *GDP_ratio* on regionalist vote depends on the distance at which the given region is from the capital of the country, and the proportion of speakers of a regional distinctive language. In two regions with no vernacular language, which are

100 km away from the capital and whose *GDP_ratio* differs by one unit, the predicted difference of regionalist support is only 2.18%. This difference becomes much higher (43.68%) for a 2000 km distance. The effect of the interaction with *language* is also very large: for two regions that have 25% of speakers of a vernacular language, the difference in expected support for regionalist parties would be 13.09% (if located at 100 km from the capital) and 54.59% (at 2000 km from the capital). With a 100% of vernacular language speakers in these regions, differences would become 45.84% and 87.34%, respectively.[21]

To sum up, from these analyses, we conclude that the following:

- The degree to which cultural differences between a region and the majoritarian culture in its country affect the support for regionalism in Western Europe is significantly influenced by relative regional affluence. Being more culturally differentiated is associated with a large gain in support for regionalism in relatively rich regions, but with a much smaller one for relatively poor ones.
- Being relatively richer only increases support for regionalism in regions whose culture is significantly different from the predominant one in the country.

Figure 1 illustrates the relevance of the interaction term for the model predictions. An increase in *GDP_ratio* translates into different predictions of votes for regionalist parties depending on the distance at which the corresponding region is from the capital of the country. While regionalist parties in a relatively rich region with a *GDP_ratio* of 0.8 and located 500 km away from the capital of the country are expected to obtain around 20% of the votes, this share is almost 60% in a region located at 2000 km from the capital, keeping the other variables at their means. Equivalently, the effect of differences in *distance* on the share of votes for regionalist parties is much
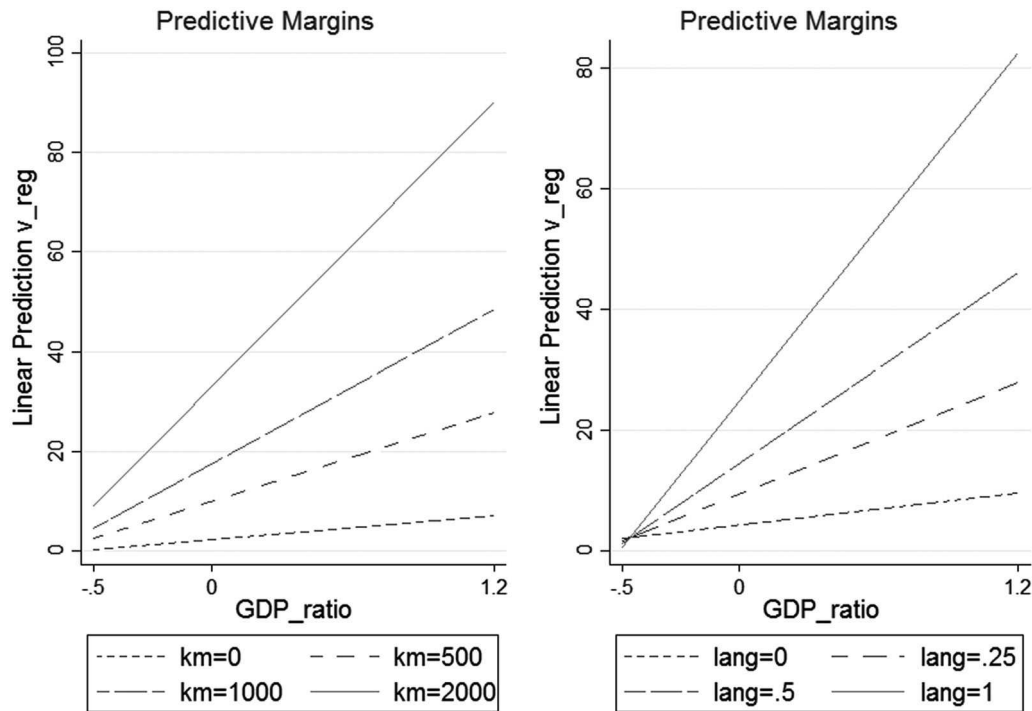
Figure 1. (left) Predicted vote for regional parties (*v_reg*) as a function of *GDP_ratio* for given distance (in 100 km); and (right) predicted vote for regional parties (*v_reg*) as a function of *GDP_ratio* for given values of language (in proportion 0–1). The other variables are evaluated at their averages.

smaller for poor regions. When we compare two equivalent regions with a *GDP_ratio* of −0.5 but at 500 and 2000 km from the capital, we see that the difference in the percentage of predicted votes is less than 10%. For the same difference in *distance*, the predicted share of votes for regionalist parties in two rich regions with a *GDP_ratio* of 0.8 differs by almost 40%.

Similarly, for a relatively rich region (*GDP_ratio* = 0.8), voting intention for regionalist parties fluctuates between 15%, 25% and 47%, as the percentage of speakers of the regional distinctive language changes from 25%, to 50% and 100%, *ceteris paribus*. In a relatively poor region (*GDP_ratio* = −0.5), the predicted support for regionalists is almost zero. In this case, having different proportions of speakers of a minority language does not affect regionalist voting.

We conducted several tests to check whether our findings are robust to different methodologies and specifications. Including only one of the two cultural variables and its interaction with *GDP_ratio* is enough to cancel out the latter's effect and obtain a positive and significant interaction (see Table A4 in Appendix A in the supplemental data online). Results hold if we use only the last regional election for every case (*N* = 120; see Table A5 in Appendix A), and if we run tobit models (see Table A7 in Appendix A) which are left- and right-censored at 0 and 100 respectively to reflect the fact that the dependent variable, by definition, can never fall outside these bounds. The interaction between cultural and economic factors is also robust to instrument *distance* and *distance_GDP* with *language* and *language_GDP* in a 2SLS regression.[22] This model should capture the exogenous effect of cultural

distance and its interaction with *GDP_ratio* (see Table A6 in Appendix A). Finally, the interactions are also robust to running a model with frequency weights, in which each region gets a weight proportional to its population (see Table A8 in Appendix A).

Returning to the main model (regression 4, Table 2), our controls report some significant effects. A 1 million increase in a region's population is associated with a decrease of a 1.65% in the expected percentage of votes for regionalist parties. This result is very robust across different specifications,[23] and also for the models with either *v_sec* or *v_auton* as dependent variables – with the separated sample.[24] Thus, having a larger weight on nationwide politics seems to ease regionalist aspirations. Coefficients associated to the variables *year*, *regional_elect*, *chamber_size* and *richer_neighbour* are positive, as expected, while having a neighbouring country which is poorer than the current one (*has_neighbour* = 1 and *richer_neighbour* = 0) decreases regionalist support – at the 10% level of significance. Including country fixed effects improves the explanatory capacity of the model, and does not have much impact on the size of coefficients.

Hypotheses 2

We proceed next to analyze whether *v_sec* and *v_reg* should be kept aggregated in *v_reg* or should be better analyzed independently. We ran a Wald test to test whether each coefficient is equal across the two models being compared. Table 4 reports the number of statistically different coefficients in each model comparison.[25] The third row shows that the Wald test rejects the hypothesis of equal coefficients for three variables: there are three coefficients

Table 4. Number of different coefficients between the models.

Models compared	Number of differences	Different coefficients
v_auton sep – *v_sec* sep	0	–
v_sec full – *v_auton* full	0	–
v_auton full – *v_auton* sep	3	*year, language_GDP, language*
v_sec full – *v_sec* sep	2	*richer_neighbour, population*

Note: Wald test, 10% level of significance. 'Full' stands for the full sample, 'sep' for the separated sample, which excludes the regions with a positive amount of the other type of regionalist vote.

which are different between the model explaining *v_auton* using the full sample, i.e., including those regions with some *v_sec*, and the model explaining *v_auton* using the separated sample, only with regions without secessionist votes. When *v_sec* becomes the dependent variable two statistically different coefficients are reported for the same models with different samples (row 4). Hence, hypothesis 2a holds: for the models using *v_sec* and *v_auton* as the dependent variables, the regression coefficients corresponding to the same variables across two different samples show statistically significant different results.

While the evidence provided so far is coherent with the argumentation for the coefficients in one of the models to be biased, it does not indicate which of the two options – full or separated sample – is better. Two pieces of evidence offer some guidance in this choice. The first – arguably limited – argument concerns explanatory capacity. The two models with the separated samples have a higher explanatory capacity than with the full samples ($R^2 = 0.52$ versus 0.35 for *v_sec* and 0.62 versus 0.45 for *v_auton*; see columns 2 and 4 of Table 5 and Table A10 in Appendix A in the supplemental data online).

The second piece of evidence is given by the result of testing hypothesis 2b. It has been argued that if both *v_sec* and *v_auton* are explained by similar factors, the coefficients of a model studying these two phenomena together using a joint sample, would be biased. Hypothesis 2b helps us in answering this question as it gives a quantifiable measure of how different are the coefficients between the models explaining *v_sec* and *v_auton*. A Wald test does not reject the null hypothesis of the coefficients being equal for any of the variables – not even at the 10% level. This holds for both models using the separated and full samples (rows 1 and 2, Table 4). Hence, hypothesis 2b is also satisfied, as there are no statistically different results between the coefficients in the models explaining the support for autonomist and secessionist parties.

In short, all available evidence supports aggregating the two subtypes of regionalist voting together. However, given that they have been frequently studied separately in the existing literature, it seems relevant to assess whether the main results for the determinants of *v_reg*, and specifically the interaction between cultural and economic factors, hold for each form of regionalism.

Hypothesis 3

Since this paper points towards an important bias in the results for *v_sec* and *v_auton* using the full sample, we will test whether the interactions between cultural and economic variables are significant with separated samples. In order to shed light on the determinants of *v_sec*, we drop the observations with a positive value for *v_auton* ($N = 759$), and conversely we drop the positive observations for *v_auton* when studying *v_sec* ($N = 822$) (Table 5).

Although *distance_GDP* loses its significance,[26] the interaction *language_GDP* keeps it, thus confirming hypothesis 3. Also, the effect of *GDP_ratio* vanishes as we bring in the interactions. These findings are robust to using tobit instead of OLS, including frequency weights, using only the last cross-section of the data, running a 2SLS model (instrumenting distance and its interaction with language and its interaction), and including only one of the cultural variables and its interaction. Some minor exceptions aside,[27] hypothesis 3 is satisfied and the interaction effect between cultural and economic variables is positive and significant in models explaining secessionism and autonomism.

Despite potential concerns regarding research design (i.e., the samples have been artificially separated), as the results for secessionist and autonomist support are consistent and comparable with those for *v_reg*, the interdependency of cultural and economic explanatory factors is reinforced.

Exploring other variables and related robustness checks

This section briefly explores the additional controls used, on top of those reported in the tables. Some of these controls are excluded because they were not significant in the main regression, or not robust to different specifications, and do not modify any result significantly.[28] The other reason to exclude some of the controls was their endogeneity problem, namely *regional_auton* and *ENERP*. These two variables have played a preeminent role in the literature so far (e.g., Massetti & Schakel, 2013; Sorens, 2005), but suffer from an important problem of simultaneity, as they are directly affected by the dependent variable. To tackle this problem, we have instrumented them with more exogenous variables.

ENERP is instrumented through *chamber_size*, a crucial aspect in the magnitude of district (Taagepera & Shugart, 1989), and a 2SLS regression is run. We find a 2SLS-positive and significant coefficient for *ENERP* – *F*-statistic test for weak instruments = 9.10 ($p = 0.0031$) – implying that a higher degree of multipartism is associated with a

Table 5. Vote for secessionist and autonomist parties.

	(1) v_sec	(2) v_sec	(3) v_auton	(4) v_auton
GDP_ratio	5.584*	–0.720	9.046**	–6.875*
	(3.023)	(5.099)	(4.267)	(3.601)
distance	0.619	0.638	0.438***	1.025***
	(0.471)	(0.481)	(0.162)	(0.294)
language	5.266*	12.89**	5.905	16.01***
	(3.024)	(5.018)	(4.133)	(4.469)
distance_GDP		0.135		1.128
		(1.434)		(0.690)
language_GDP		40.33***		63.66***
		(13.83)		(13.76)
year	0.102***	0.103***	0.0957***	0.111***
	(0.0287)	(0.0277)	(0.0316)	(0.0320)
has_neighbour	0.664	–1.060	1.102	–2.003
	(1.338)	(1.396)	(1.368)	(1.235)
richer_neighbour	5.842**	5.653**	5.791**	5.104***
	(2.685)	(2.663)	(2.430)	(1.839)
population	–0.115**	–0.118**	–0.157**	–0.130***
	(0.0497)	(0.0468)	(0.0610)	(0.0433)
chamber_size	0.0804**	0.0906**	0.0778*	0.0900***
	(0.0404)	(0.0378)	(0.0395)	(0.0343)
regional_elect	1.279	1.254	1.226***	1.084***
	(0.793)	(0.777)	(0.384)	(0.366)
Country fixed effects	Yes	Yes	Yes	Yes
Observations	759	759	822	822
R^2	0.492	0.519	0.459	0.619
Adjusted R^2	0.480	0.506	0.447	0.609

Notes: Pooled ordinary least squares (OLS) model explaining votes for secessionist (columns 1 and 2) and autonomist parties (columns 3 and 4). The regression for each type of regionalist vote excludes those regions with a positive amount of the other type of regionalist vote (separate samples). Errors are clustered at the regional level. The constant term is excluded for paucity. Standard errors are given in parentheses.
*$p < 0.10$, **$p < 0.05$, ***$p < 0.01$.

better electoral performance of regionalist parties (see column 1 in Table A9 in Appendix A in the supplemental data online). Similarly, we instrument *regional_auton* with *hist_together*. While the F-statistic of excluded instruments for the first stage is a bit low (2.65, $p = 0.106$), the 2SLS coefficient for *region_auton* is 29.17, and significant at the 1% (see column 2 in Table A9). This points towards higher regional autonomy having a positive effect on regionalist support.

There are two additional caveats to our results: (1) the main model does not take into account irredentist particularities; and (2) some parties are present in several regions at the same time, which might make these observations non-independent.

When we incorporate *irredentist_potential* into our model. It does not have an impact on regionalist voting and does not change our findings (column 4, Table 2). Additionally, when we drop the 18 territorial units where this variable was positive, our main results still hold (see column 3 in Table A9 in Appendix A in the supplemental

data online). To control for potential problems derived from the same regionalist parties contesting elections in different regions, we include the *party_elsewhere* variable. Again, when we include this variable in the regression, it is not significant and does not affect the results. These are also robust to dropping the observations corresponding to the 17 regions with relevant regional parties contending in some other region (e.g., Lega Nord, PNV, etc.) (see column 4 in Table A9 in Appendix A).

CONCLUSIONS

This paper analyzes the effect of cultural and economic factors on voters' desire for greater regional autonomy and independence using a panel of territorial subunits in 10 West European democracies. We extended the traditional theory using the concept of cultural proximity. The main argument is that the population of richer regions is more likely to tolerate transfers to poorer regions as long as they feel part of the same national group. In relatively

rich and culturally differentiated regions, there should be a larger political support for regionalism. This hypothesis is tested empirically using a dataset with the results from regional elections in these countries between 1969 and 2000 – updated with the most recent elections till 2014. Evidence in support of the predictions of the theoretical framework is found in the form of statistically significant interaction effects between cultural and economic variables.

Moreover, regional relative affluence only has a positive effect on the support for regionalism through these interaction terms: a region being relatively richer is only associated with an expected higher support for regionalist parties to the extent that it is culturally differentiated (i.e., the proportion of minority language speakers and physical distance are larger). Equivalently, the degree to which a region's cultural difference is associated with a higher support for regionalist parties depends on its relative income, as this effect is much larger for rich than for poor regions. Also, both theory and empirical evidence suggest that support for autonomist and secessionist parties should be considered together in electoral studies: the interaction between income and culture underpins both, and there are good reasons to think that models using common samples to study these two phenomena separately will be biased.

Further research is necessary to unravel plausible interaction effects between income and cultural variables beyond Western Europe, and also to understand why the interactions do not hold for some particular cases. Additional theoretical and empirical studies that incorporate microdata should complement this study by analyzing whether this interaction effect applies at an individual voter level – and, in that case, exploring the concrete mechanism that might explain it. Also, locating all (regionalist and non-regionalist parties) that compete in an election on a territorial continuum that ranges from extreme secessionism to maximum unionism (e.g., on the basis of party manifestos per each election) might help us to understand further how this interaction operates.

To summarize, the main theses of the paper are supported by our empirical analyses. Firstly, cultural and economic factors affect regionalist voting through a positive and significant interaction – this even makes the relative affluence of the region insignificant on its own. Secondly, keeping aggregated autonomist and secessionist electoral support in a unique regionalist category – instead of considering them as two separated phenomena – seems more adequate. Finally, the significant interaction between income and culture in the separate explanation of autonomist and secessionist support reinforces its relevance in any electoral study of minority nationalism.

NOTES

1. Nation-states combine the state as a system of political action with national identity, the community's entitlement to claim for self-rule (Keating, 2001).
2. In the vast majority of cases, secessionism demands take the form of separatism, oriented to supplying regions with independent state structures. There are exceptions, however, as people in some regions would rather prefer to join another state. For instance, the German-speaking community in the Italian South Tyrol may prefer a reattachment to Austria instead of forming a state on their own. This phenomenon is known as *irredentism*, which is frequently considered a specific subtype of secessionism (Sorens, 2005, p. 308, 2008, pp. 339–340).
3. A 'region' is a 'territorial body of public law established at the level immediately below that of the state and endowed with political self-government' (Assembly of European Regions, 1996, Art. 1.1, p. 4; see also Hooghe et al., 2010). We use 'regions' and 'territorial sub-units' interchangeably throughout. A 'nation' is a community of people with a set of coordinated beliefs about their cultural identities who claim self-government, with varying degrees of intensity (Laitin, 2007, pp. 40–41).
4. 'National identity' is the sense of a nation as a cohesive whole. It refers to the set of tendencies and values derived from the feeling of belonging to a united group of people with distinctive traditions, history and culture. National identity is neither fixed nor alterable at will. It requires a periodical redefinition in the light of historical features, present needs and future aspirations (Parekh, 1995).
5. Considering the following parties as secessionists: CiU, ERC, ICV, CUP and SI in Catalonia; and PNV and Bildu in the Basque Country (Nordsieck, 2014) (see Table A2 in Appendix A in the supplemental data online).
6. The Basque language is also spoken in Navarre, where there is indeed a relatively high support for both autonomist and secessionist parties.
7. Munoz and Tormos (2014) emphasize the importance of economic aspirations and partisanship in their enquiry on individual-level determinants for secessionism in Catalonia.

8. We exclude Dutch, British and French overseas countries and territories (and also the Channels Islands and Isle of Man). The same holds for Greenland (Denmark), given its quasi-independent status. In turn, we include outermost regions (Azores, Madeira, Canary Islands, French overseas departments). We also consider the following 'special cases': Ceuta and Melilla (Spain), Gibraltar (UK) and Faroe Islands (Denmark).

9. Our data for the percentage of votes casted for a party come both from Nordsieck's (2014) *Parties and Elections* database and official regional or state-level resources. When more than one party fits in the secessionist or autonomist category, we aggregate their percentages. To be included, the party must have gained either seats in the regional chamber or more than 1% of votes.

10. Following Massetti and Schakel's (2013, p. 11) criteria, in electoral coalitions where regionalist parties are involved, seat allocation is used to assign vote shares to the coalition partners.

11. In some cases, it can be challenging to classify political parties. For example, it is not clear to which extent the Bavarian CSU, Northern Irish UUP and Navarra's UPN act separately from statewide parties (Spinner-Haleve & Theiss-Morse, 2013, pp. 11–12). We rely on Nordsieck (2014) and Sorens (2005) criteria, and considered the party as regionalist or not depending on the election. Excluding these problematic cases does not change our findings.

12. This refers to long-term established regional languages, excluding recent linguistic minorities (e.g., first generations of migrants).

13. Estimating the proportion of speakers and defining the linguistic status (i.e., whether it is a dialect or a language) is often problematic. Although *Language* is time-invariant, it should be a good proxy for cultural heterogeneity.

14. The correlation holds if the analysis by Desmet et al. (2011) is restricted to the countries in our sample (0.44). It is always higher than the correlation between linguistic and cultural distance: 0.42 in their sample, 0.33 for ours.

15. We run a pooled ordinary least squares (OLS) regression with Sorens (2005) panel subsample for these three countries ($N = 1000$). As expected, while *distance* and the other related variables are never significant, *language* is still significant.

16. The hypothesis of weak instrument is rejected (in model 4, the F-test for excluded instruments is $F(2,120) = 6.23$, $p = 0.0027$, for *distance*; and 4.08, $p = 0.0194$, for *distance_GDP*).

17. Given that no observation in our sample gets close to a zero value for *GDP_ratio*, centring this variable around 1 is more appropriate (the lowest value in our sample is 0.4; observations below 0.6 are rare).

18. The earliest election in our dataset is in 1969, the last in 2014. Data in this full interval of time are not available for all regions and elections. Even though historical and institutional differences between the countries might be important, given that the effect of country-level factors on the support for regionalism is not the focus of our study, we include country fixed effects.

19. Beyond the scope of this paper, this finding presents an interesting avenue for enquiry for the social psychological literature on national identity (e.g., Bloom, 1993; Spinner-Haleve & Theiss-Morse, 2003).

20. We explored the implications of the region having the capital of the country in an unreported regression with a dummy. Its coefficient is not statistically significant once geographical distance is controlled for, and does not alter the results.

21. This is an extreme case, but in fact more than 90% of the population speaks a vernacular language in some sub-units, such as the German-speaking communities in Belgium or the Faroe Islands in Denmark.

22. In this specification, the coefficient for *distance_GDP* is larger than in the main model. The effect of *GDP_ratio* on *v_reg* becoming negative and significant for the 2SLS regression means that the population in not culturally differentiated regions which are richer would vote less for regionalist parties than the population in poorer ones, taking the main thesis of the paper to the extreme.

23. The negative significant coefficient for *population* holds also in non-reported robustness checks, such as including the country's population, or the relative population of the region over the population of the country as regressors, or taking the log of the variable instead of the level.

24. Population loses its significance and changes its sign only in the regressions with the full sample and *v_sec* as the dependent variable (see columns 1 and 2 in Table A10 in Appendix A in the supplemental data online). We contend throughout that these models are biased.

25. The explanatory variables included in each of these models are the same as those included in the main model (see column 4 in Table 2). The models in Table 4 are respectively: *v_sec* sep (see column 2 in Table 5); *v_auton* sep (see column 4 in Table 5); *v_sec* full (see column 2 in Table A10 in Appendix A in the supplemental data online); and *v_auton* full (see column 4 in Table A10).

26. The p-value for *distance_GDP* is 0.105, though (see column 4 in Table 5).

27. There are two exceptions. With *v_sec* as the dependent variable, if only *distance* and its interaction are included as cultural variables, neither *GDP_ratio* nor the *GDP_distance* interaction are significant, and if we use only the last cross-section, the interactions are not significant ($N = 75$).

28. These include *distance_sq*, *EU*, *religion*, *indep* and *capital*, as well as the legal status of the language (whether or not if it is official), whether or not the region is on the mainland, population density, the population of the country, the proportional population of the region over that of the country, and the ideological gap.

REFERENCES

Alesina, A., & Spolaore, E. (1997). On the number and size of nations. *Quarterly Journal of Economics*, 112(4), 1027–1056. doi:10.1162/003355300555411

Assembly of European Regions. (1996). *Declaration on regionalism in Europe*. Assembly of European Regions. Basel: Assembly of European Regions. Retrieved from http://aer.eu/aer-declaration-regionalism/

Bloom, W. (1993). *Personal identity, national identity and international relations* (Cambridge Studies in International Relations). Cambridge: Cambridge University Press.

Bolton, P., & Roland, G. (1996). Distributional conflicts, factor mobility and political integration. *American Economic Review*, *86*(2), 99–104.

Bolton, P., & Roland, G. (1997). The breakup of nations: A political economy analysis. *Quarterly Journal of Economics*, *112*(4), 1057–1090. doi:10.1162/003355300555420

Brancati, D. (2014). Another great illusion: The advancement of separatism through economic integration. *Political Science Research and Methods*, *2*(1), 69–95. doi:10.1017/psrm.2013.30

de Winter, L. (1998). Conclusion: A comparative analysis of the electoral, office and policy success of ethnoregionalist parties. In L. de Winter & H. Türsan (Eds.), *Regionalist parties in Western Europe* (pp. 204–247). New York: Routledge.

Della Porta, D., O'Connor, F., Portos, M., & Subirats, A. (2017). *Referendums from below: Social movements and direct democracy in the neoliberal crisis*. Bristol: Policy Press.

Desmet, K., Le Breton, M., Ortuño-Ortín, I., & Weber, S. (2011). The stability and breakup of nations: A quantitative analysis. *Journal of Economic Growth*, *16*, 183–213. doi:10.1007/s10887-011-9068-z

Ethnologue. (2013). *Ethnologue: Languages of the world 2013*. Retrieved September 6, 2015, from https://www.ethnologue.com/.

Fearon, J. D., & van Houten, P. (2002). *The politicization of cultural and economic difference. A return to the theory of regional autonomy movements*. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.476.1040&rep=rep1&type=pdf

Gordin, J. P. (2001). The electoral fate of ethnoregionalist parties in Western Europe: A Boolean test of extant explanations. *Scandinavian Political Studies*, *24*(2), 149–170. doi:10.1111/1467-9477.00051

Hale, H. E. (2008). *The foundations of ethnic politics: Separatism of states and nations in Eurasia and the world*. Cambridge: Cambridge University Press.

Hearl, D. J., Budge, I., & Pearson, B. (1996). Distinctiveness of regional voting: A comparative analysis across the European Community countries (1979–93). *Electoral Studies*, *15*, 167–182.

Hooghe, L., Marks, G., & Schakel, A. H. (2010). *The rise of regional authority: A comparative study of 42 democracies (1950–2006)*. London: Routledge.

Horowitz, D. (1981). Patterns of ethnic separatism. *Comparative Studies in Society and History*, *23*, 165–195. doi:10.1017/S0010417500013268

Jolly, S. K. (2006). *A Europe of regions? Regional integration, subnational mobilization and the optimal size of states* (PhD dissertation). Duke University, Durham, NC.

Keating, M. (2001). *Nations against the state: The new politics of nationalism in Quebec, Catalonia and Scotland*, 2nd ed. Houndmills: Palgrave Macmillan.

Laitin, D. D. (2007). *Nations, states and violence*. Oxford: Oxford University Press.

Massetti, E. (2009). Explaining regionalist party positioning in a multi-dimensional ideological space: A framework for analysis. *Regional and Federal Studies*, *19*(4/5), 501–531. doi:10.1080/13597560903310246

Massetti, E., & Schakel, A. (2013). Between autonomy and secession: Decentralization and regionalist party ideological radicalism. *Party Politics*. doi:10.1177/1354068813511380

Medeiros, M., Gauvin, J.-P., & Chhim, C. (2015). Refining vote choice in an ethno-regionalist context: Three-dimensional ideological voting in Catalonia and Quebec. *Electoral Studies*, *40*, 14–22. doi:10.1016/j.electstud.2015.06.004

Munoz, J., & Tormos, R. (2014). Economic expectations and support for secession in Catalonia: Between causality and rationalization. *European Political Science Review*. doi:10.1017/S1755773914000174

Nordsieck, W. (2014). *Parties and elections in Europe*. Retrieved October 19, 2014, from http://www.parties-and-elections.eu/countries.html

Parekh, B. (1995). The concept of national identity. *Journal of Ethnic and Migration Studies*, *21*(2), 255–268. doi:10.1080/1369183X.1995.9976489

Riera, P. (2014). *Nationalist parties and voting in the Balearic Islands* (Unpublished manuscript). Retrieved from http://recp.es/uploads/files/modules/congress/12/papers/1373.pdf

Sorens, J. (2004). Globalization, secessionism, and autonomy. *Electoral Studies*, *23*, 727–752. doi:10.1016/j.electstud.2003.10.003

Sorens, J. (2005). The cross-sectional determinants of secessionism in advanced democracies. *Comparative Political Studies*, *38*(3), 304–326. doi:10.1177/0010414004272538

Sorens, J. (2008). Regionalists against secession: The political economy of territory in advanced democracies. *Nationalism and Ethnic Politics*, *14*(3), 325–360. doi:10.1080/13537110802289670

Sorens, J. (2012). *Secessionism: Identity, interest, and strategy*. Montreal: McGill-Queen's University Press.

Spinner-Haleve, J., & Theiss-Morse, E. (2003). National identity and self-esteem. *Perspectives on Politics*, *1*(2), 515–532. doi:10.1017/S1537592703000379

Taagepera, R., & Shugart, M. S. (1989). *Seats and votes: The effects and determinants of electoral systems*. New Haven and London: Yale University Press.

Van Houten, P. (2000). *Regional assertiveness in Western Europe. Political constraints and the role of party competition* (PhD dissertation). University of Chicago, Chicago.