

The Cross-Sectional Determinants of Secessionism

Jason Sorens¹

Yale University, Department of Political Science

May 2002

Abstract. This paper analyzes the “risk factors” of secessionism at the substate, regional level. It seeks to answer the question, What regions are more likely to support more successful secessionist parties? Using new data in pooled, cross-sectional time-series regression analysis, I find that the following factors best explain secessionist vote share: regional language, non-contiguity of region with independent country in which regional language is an official language, history of independence, non-contiguity of region with rest of country, relative affluence of region, absolute size of region in terms of population, and multi-party political system. Some of these variables have an effect on secessionism conditional on pre-existing linguistic difference.

¹ The author wishes to thank Geoffrey Garrett, Nathan Jensen, Fiona McGillivray, Frances Rosenbluth, and Jose Cheibub for helpful comments. The author also thanks the staff of the Statistical Library of Denmark and of the University of California at San Diego library system for their assistance.

The scholarly study of secessionism has gone in cycles, lagging real-world trends. In the 1970s and early 1980s, studies of “ethnonationalism,” “minority nationalism,” “micro-nationalism,” and the like exploded, following the remarkable electoral successes of the nationalists of Scotland and Quebec.¹ This literature was preoccupied with explaining why the literature of the 1950s and 1960s predicting the inevitable demise of peripheral nationalism under the pressures of modernization turned out to be wrong. However, following the electoral setbacks suffered by the Scottish National Party (SNP), Plaid Cymru/Party of Wales (PC), and Parti Québécois/Quebecker Party (PQ) in the late 1970s and early 1980s, another wave of revisionism eventually set in, explaining how secessionist leaders could be co-opted and their movements deflated.² In the mid and late 1980s armed secessionist conflict in Africa and Asia brought the issue back to the foreground for students of ethnic conflict in the Third World. The breakup of the Soviet Union, Czechoslovakia, and Yugoslavia provided more material for explanation in the 1990s.³ The comparative study of secessionism in well-established democracies has not, however, picked up where it left off in the 1970s, despite the persistence and even growth of secessionism in Scotland, Wales, Euskadi (the Basque Country), Catalonia, Corsica, and Flanders, the 1995 Quebec referendum on secession, the meteoric rise and fall of the Lega Nord/Northern League (LN) in Italy, the imminent Faroese

bid for independence, and the emergence of small secessionist parties in most advanced industrial democracies.⁴ As a consequence, political scientists know very little about the determinants of secessionism across regions in democracies.

This paper formulates hypotheses that can be applied across regions in all well-established democracies and tests them using a pooled, cross-sectional time-series regression format. To my knowledge, this study is the first of its kind. Ayres and Saideman have in two articles explored the determinants of secessionism and irridentism over a global sample.⁵ There are, however, several limitations of their approach: 1) a dummy dependent variable (coded 1 when a group is separatist) does not allow for the model to predict varying degrees of secessionism; 2) a global sample can be inappropriate when the effects of certain variables change depending on some other factor such as regime type, a possibility explored below; 3) their study focuses on ethnic groups rather than regions – however, some secessionist movements (in Scotland, Alaska, Savoy, and Padania,⁶ for example) are not primarily ethnically based; 4) in their study many of the independent variables had to be coded at the state level, rather than the sub-state group or regional level, which is the preferred unit of analysis.

This paper addresses these issues by limiting the sample to well-established democracies where secessionist parties are legal (not

India and Turkey, for example) and where significant regional differentiation exists (not Iceland and Jamaica, for example),⁷ using secessionist party vote share to construct the dependent variable, using existing statutory regions/provinces as the units of analysis, and using a new dataset with cultural, demographic, political, and economic data disaggregated to the provincial level. Restricting the sample to democracies has theoretical justification, since some of the independent variables should have different effects in autocracies than in democracies, but it also has a major empirical benefit. It allows the use of a continuous variable with low measurement error, vote share, for the dependent variable. To the extent that secessionist vote share does not perfectly track popular support for secession, these deviations are fairly regular and predictable: they can be controlled for in the model through use of a lagged vote share variable and a variable for multi-partism, for example.

Hypotheses

What factors should explain secessionist party vote share across provinces? The most obvious base model would be that secessionism correlates with the existence of a minority language. A region-specific language generally captures the “ethnic” component of secessionism, though in some areas religion may be a better marker of ethnicity than language (Northern Ireland, for example).

Hypothesis 1: A region-specific minority language increases secessionism.

To test this hypothesis, a variable for percentage of provincial population speaking a minority language peculiar to that territory (and in some cases, surrounding areas)⁸ at the latest available date is used: *LANG*. The variable's value at the latest date is carried back in time for all values for a panel (province), to make sure that a decline in language speakers does not confound the results. While we might reasonably expect that a greater number of speakers of a minority language across regions correlates with greater secessionist vote, the same correlation does not necessarily hold true over time. In fact, a precipitous decline in speakers of a minority language might actually stimulate secessionism as a means of protecting the language from extinction.

Not all region-specific minority languages are equal when it comes to secessionism, however. A language spoken in a minority enclave in one country might be the dominant language of a nearby, sizeable country. In this case, irridentism - support for a policy of separating the territory from one state and attaching it to another - seems a more likely outcome than secessionism. Thus, German speakers in South Tyrol (officially, Bolzano) are more likely to seek re-attachment to Austria than an independent South Tyrol. In practice, cases of potential irridentist conflict are usually alleviated by

negotiations between the “big countries” involved, who wish to avoid persistent international tension over the issue. Thus, these potentially irridentist enclaves typically receive high levels of autonomy and “special status” within their state.⁹ Another reason why such enclaves might be less secessionist than regions without adjoining states controlled by their language group is that cross-border cultural exchange should make the language group feel less threatened. Minority language groups without such ties, such as the Welsh and Basques, realize that if the language fails in their region, it dies out completely.¹⁰

Hypothesis 2: Linguistically distinctive regions that are contiguous to independent countries in which the regional language is an official language are less likely to support secessionist parties.

To test this hypothesis a dummy variable, *LANGADJ*, is included in the regressions: it is coded “1” for provinces with a region-specific minority language and which are contiguous to independent countries in which that language has official status and “0” otherwise.

Some secessionist movements are considered to have a “civic” rather than “ethnic” basis for their nationalism. This civic identity might derive from a history of independence. For example, Scottish nationalism is often said to be based on the extensive Scottish history

of independence, while Welsh nationalism is thought to be based on defense of the Welsh language.

Hypothesis 3: Regions with a recent history of independence are more likely to support secessionist parties.

To test this hypothesis, a dummy variable called *INDEP* is employed. The variable is coded “1” when a province has been an independent country (by its own lights) or served as the capital province of an independent country different from the current state at any time since 1648, “0” otherwise. The date 1648, the year of the Treaty of Westphalia, is used because it is often taken to signify the end of the feudal system of family properties and the beginning of modern state nationalism.

Secessionist demands are not always based on traditional conceptions of identity. In some notable historical cases, the basis for secession has been economic.¹¹ Territorial grievances are frequently based on perceptions that the region is a net “loser” from the existing political union. In particular, when a region pays more in taxes than it receives in expenditures it is likely to present demands for fiscal autonomy and possibly even secession if those demands are not met. Regions that receive more in expenditures than they pay in taxes are poor ground for secessionism, because independence would mean the loss of subsidies.

Hypothesis 4a: Regions that pay more in taxes to the central government than they receive back in

expenditures are more likely to support secessionist parties.

To test this hypothesis, the *GDPRATIO* variable is used. This variable measures the per capita GDP of the province divided by the per capita GDP of the country. In democratic countries, more affluent regions are more likely to be losers from fiscal union, because under democracy redistribution from wealthy to poor citizens is a typical outcome. Regions with poorer citizens should benefit from this state of affairs. Thus, a corollary of Hypothesis 4a would be:

Hypothesis 4b: In democracies, relatively affluent regions are more likely to support secessionist

parties than relatively poor regions.

This hypothesis directly contradicts Horowitz's observation that poor groups in poor regions are early, frequent secessionists, while advanced groups in advanced regions are late, rare secessionists.¹²

Horowitz's hypothesis appears to hold up well in autocracies, however, which constitute the bulk of the cases from which he draws his conclusions. The reason the hypothesis holds in autocracies is likely that in autocracies access to the resources of the state is determined by relative regional affluence: regional redistribution in autocracies tends to be regressive, unlike in democracies. Poorer regions in autocracies fear the domination of more affluent regions

and are thus more likely to resist tight union. In democracies, poorer regions have the upper hand, and wealthier regions fear tight union.

Not all economic discrimination can be measured in terms of fiscal balance, however. The state can also discriminate against certain regions in its economic management. Presumably this phenomenon underlies the charge of “internal colonialism,” the idea being that the central government exploits and represses the growth of peripheral regions for the benefit of the core. It has been discovered that increases in provincial unemployment relative to countrywide unemployment correlate with increases in secessionist vote.¹³ The “short panels” in the regressions in this paper mean that time-series effects on secessionism are unlikely to appear clearly, however. The large number of panels means that cross-sectional effects should show up more clearly than time-series effects: a variable that explains differences across provinces well will show up as more important in the results than a variable that explains changes over time well. Nevertheless, it is important to include these time-series variables as controls.

Hypothesis 5: When regional unemployment rises relative to the rest of the country, secessionist vote rises.

The regressions thus include a variable measuring provincial unemployment rate minus countrywide unemployment rate:

UNEMDIFF.

Absolute unemployment rate has different effects, however. It has been likewise discovered that increases in a provincial misery index correlate with decreases in secessionist vote.¹⁴ The reason is presumably that hard economic times move voters' focus from issues of constitutional change to bread-and-butter economic issues, which are usually fought on the traditional left-right dimension, a fact that benefits traditional parties of left and right.

Hypothesis 6: When a region's economy does poorly, secessionist vote falls.

Again, this hypothesis is a time-series one and results may not show up well in this paper's regressions. However, a variable measuring provincial unemployment rate plus country inflation rate (*MISIND*) is used as a control.

A final time-series hypothesis is tested in this chapter: that globalization increases secessionism. The argument is that globalization reduces the economic inefficiency of secession by allowing small territories to exploit their comparative advantage without being part of a larger political unit. (For a survey of the literature see Sorens (2002).)

Hypothesis 7: Globalization increases secessionism in all regions.

As a control variable, then, a variable measuring world exports divided by world production is used: *GLOB*.¹⁵ The variable is rebased as an index with the 1990 value set to 1. For any given year, the variable takes the same value for all provinces.

One demographic-economic variable that should have an observable cross-sectional effect is population. Larger territories are more viable as independent states than small ones.¹⁶

Hypothesis 8: More populous regions should have higher secessionist vote.

The *POP* variable measures provincial population.

Besides cultural and economic variables, political institutions also matter for secessionism. Horowitz argues that secessionism is more likely where territorial units within a country are few and large.¹⁷ In these cases there are only a few vectors of conflict, and compromise becomes more difficult. Also, when an ethnic group is unified in a single territory, collective action against the state is easier. It is no accident that the most centralized states in the world (e.g., Turkey, the Baltic states, Greece) tend to have many very small provinces rather than a few large ones.

Hypothesis 9: Provinces that are larger relative to the state as a whole are more likely to support

secessionist movements.

To test this hypothesis, a variable *REP* is employed, measuring provincial representation in the lower house of the central state's parliament as a percentage of total seats in the house.¹⁸

Two other political institutions are potentially important: regional autonomy and multi-partism. Central governments often offer regional autonomy to combat electoral secessionism.¹⁹ On the other hand, autonomy arrangements may promote a sense of distinctiveness and confidence in self-government over time. Thus, we are left with two distinct hypotheses that are in some tension with each other.

Hypothesis 10: Regions that enjoy substantial autonomy are more likely to develop secessionist parties in the long run.

Hypothesis 11: Offers of regional autonomy reduce secessionism by making a region's voters more content with the political union.

This paper attempts to parse these hypotheses with two variables: *PROVAUTO* and *DAUTO*. *PROVAUTO* measures the level of provincial autonomy, while *DAUTO* measures the change in *PROVAUTO* since the last election. The author coded the *PROVAUTO* variable in the following way:

+1 if the province has an elected executive;

+1 if the province is legally superior to geographically lower-level territorial units,

such as municipalities (that is, it may create, alter, or abolish them);

+1 if the province enjoys legislative powers in addition to administrative powers

(that is, it may assume powers not expressly delegated to it by the central

government);

+1 if the provincial government derives more than 25% of its operating budget from

own-source revenues (the provincial government having control over both rates and

types of taxation).

PROVAUTO thus takes into account political, legislative, and fiscal aspects of autonomy arrangements. It ranges from 0 to 4. American states and German *länder* are examples of provinces given a “4” score, while French *régions* were given a “0” score before the regionalization of 1982-1986.

The second important “institution,” multi-partism, is more a phenomenon that reflects the intersection of institutions and social cleavages. The number of parties in a given area is a function of the electoral rules and the dimensions of political conflict.²⁰ Where

electoral rules are disproportional, or where social practice generally discriminates against smaller parties, the number of political parties will be smaller. Where the number of political parties is smaller, secessionist vote should generally be lower, because in most areas, secessionist parties compete with at least two dominant parties reflecting the traditional left-right divide on socioeconomic issues.

Hypothesis 12: In regions where multi-partism is lower, secessionist vote will be lower.

To test this hypothesis, a variable *ENEP*, standing for “effective number of electoral parties,” is used. This variable is calculated at the provincial level from the last election of the same type, countrywide or provincial²¹ (the purpose of the lag is to try to avoid endogeneity problems – secessionist vote is reflected in this variable at the same election) using the formula developed by Laakso and Taagepera:²²

$$ENEP=1/\sum(p_i^2)$$

where p is the proportion of the vote for each party i .

Control Variables

The regression analyses employ some control variables. Since the regressions include both countrywide and provincial elections (with vote share measured at the provincial level in both cases), it is important to include a dummy variable for provincial elections.

PROVELEC is coded “1” when the election is a provincial election, “0”

when it is a countrywide election. Secessionist parties may do better in provincial than countrywide elections because they have a better chance of participating in government and thus being able to distribute patronage or provide a policy alternative to the state-wide parties. *PROVELEC* is not used when the lagged vote variable is used in regressions (see the explanation of this variable below), the reason being that the lagged vote variable is constructed from vote share in the previous provincial election for provincial-election observations. The lagged vote variable thus explains the variance between provincial and countrywide elections adequately.

Two ideological controls used in the regressions are *IDEOL* and *LEFT*. *IDEOL* measures the absolute value of provincial vote for right parties minus country vote for right parties, from the previous election. It thus is meant to capture absolute ideological differences between the province and the country. *LEFT* is simply *IDEOL* without the absolute value term: it thus measures how left-wing (or non-right-wing) a province is relative to the rest of the country.

DIST is a dummy variable measuring whether the province is connected to the rest of the country by roads or not. It is coded "1" if the province is detached from the rest of the country, "0" if it is contiguous. Prior to running the regressions, the author did not believe this variable would be important, but it turned out to be highly significant in all regressions. The conclusion of the paper speculates

on what this variable could be measuring and why it is important for secessionism.

Dependent Variable

The dependent variable is vote share for secessionist parties in elections to the lower houses of countrywide and provincial legislatures. However, it has been calculated in two different ways and then logistically transformed. The first way of calculating the variable was to include votes only for independentist parties if they existed, and if none did, vote for “moderate secessionist” or “radical autonomist” parties was used. The difference between independentism and radical autonomism can be illuminated with a few examples. The SNP are independentists: they favor the creation of an independent Scotland within the European Union. The PC is not independentist: they favor an “associated state” status for Wales that includes a seat in the United Nations. Some PC members favor independence, but the party is officially agnostic on the matter. The PC is thus a “moderate secessionist” or “radical autonomist” party. Likewise, in Euskadi Herri Batasuna/Euskal Herritarrok (HB) and Eusko Alkartasuna (EA) favor independence, while the Basque National Party (PNV) and the Basque Left (EE) acknowledge a *right* to independence but believe that full independence should not be pursued as yet. HB and EA are thus independentists, while the PNV and EE are “moderate secessionists” and are not included in the first

method of calculating the dependent variable. The dependent variable calculated according to this method was called *VOTE*.

VOTE2 was calculated in a different way: simply adding up all the votes for moderate and radical secessionists in the province. Neither *VOTE* nor *VOTE2* includes vote share for regionalist parties that favor only status-quo autonomy arrangements or that explicitly oppose independence.

VOTE and *VOTE2* were logarithmically transformed in order to smooth out their significant skewness somewhat. New dependent variables were created as follows:

$$LNVOTE = \ln(3 + VOTE)$$

$$LNVOTE2 = \ln(3 + VOTE2)$$

Three was added in order to ensure that the lowest values of the variables were at least 1. The correlation between *LNVOTE* and *LNVOTE2* is 0.95 in the full provinces dataset.

For some regressions lagged vote variables were included to account for the effect of partisan attachment. For a given secessionist vote share at election at time x , the lagged variable measures the vote those parties received at time $x-1$.²³ The lagged vote variable for *VOTE* is called *LAGVOTE*, while the lagged vote variable for *VOTE2* is called *LAGVOTE2*. The variables are then logarithmically transformed as in the above equations to create *LNLAG* and *LNLAG2*, respectively. These lagged vote variables are not quite the same as a lagged

dependent variable. If a party contests an election at $x-1$ but disappears before the election at x , its vote share at $x-1$ is not included in the lagged vote for the observation at x , unless no other secessionist party contests the election at x . For most provinces, there would be no distinction between the lagged vote variables and a lagged dependent variable, but in cases where secessionist parties replace each other frequently, such as Sardinia, there can be significant differences.

Methodological Strategies

Separate regressions were run with both of the dependent variables described above. The differences in the results are minor, however. Only the results for regressions using the *LNVOTE* dependent variable are reported in the text; see the Appendix for results using *LNVOTE2*. Regressions were also run with and without lagged vote variables. The interpretation of the results for the different sets of regressions is slightly different. For a regression in which a lagged vote variable is included, what the other independent variables are explaining is not the absolute level of secessionist vote share, but the difference in secessionist vote share between the present and immediately prior elections. The differences in results between these two sets of regressions turned out to be mostly minor, however, indicating that in general, the factors that explain absolute

secessionist vote across provinces also explain increase in secessionist vote across provinces.

While statutory first-tier territorial units are the panel units for the regressions reported in the text, a potential difficulty arises from the fact that provincial boundaries do not always accord with the boundaries of historical regions, the units for which secession would presumably be sought. Regressions were thus also run with “historical regions” as the panels. Differences in the results were usually minor, but the results for the historical-regions regressions, along with definitions of the historical regions, are reported in the Appendix.

The presence of censoring on the dependent variables has necessitated the use of a special maximum likelihood procedure, Tobit. Tobit is the standard regression procedure for datasets in which significant censoring exists. It is especially useful for datasets in which vote share is the dependent variable, since vote share cannot go below 0. (With the logistic transformation, the censoring point on *LNVOTE* and *LNVOTE2* is 1.1.) Thus, Tobit has been used before in quantitative studies of vote share.²⁴

A final issue for the regressions is the distinction among the hypotheses in terms of direct and indirect expected effects. Factors like language and history of independence should have mostly a direct effect on secessionism. Political institutions and economic conditions

should have indirect effects, however. For example, a large population will not in itself bring about the emergence of a secessionist party in a territory, nor will a multi-party political system. These variables are expected to have more importance in a dataset limited to provinces that already have the “base conditions” for secessionism. Thus, two stages of regressions are run. In the first stage, all data points are used and all variables are put in together. In the second stage, only provinces which score above zero on the *LANG* variable (that is, “linguistically distinctive” provinces) are included in the regression. Variables like *POP* and *ENEP* should show clearer results in this second stage. It is an open question whether variables like *GDP**RATIO* have a greater direct or indirect effect. Is fiscal discrimination a substitute for linguistic distinctiveness in stimulating secessionism, or do the two factors work together?

Table 1 presents the different variations of regression analysis run. The regressions for which results are presented in the text are in bold. The other results are presented in the Appendix.

Table 1

Dataset	<i>LNVOTE</i> dependent variable	<i>LNVOTE2</i> dependent variable
Provinces	All provinces; no lagged vote	All provinces; no lagged vote
Provinces	All provinces; lagged vote	All provinces; lagged vote
Provinces	Linguistically distinctive provinces; no lagged vote	Linguistically distinctive provinces; no lagged vote
Provinces	Linguistically distinctive provinces; lagged vote	Linguistically distinctive provinces; lagged vote
Regions	All regions; no lagged vote	All regions; no lagged vote
Regions	All regions; lagged vote	All regions; lagged vote
Regions	Linguistically distinctive regions; no lagged vote	Linguistically distinctive regions; no lagged vote
Regions	Linguistically distinctive regions; lagged	Linguistically distinctive regions; lagged

	vote	vote
--	------	------

Results

Tables 2 and 3 present the results from the regressions on the dataset of all provinces, without and with a lagged vote variable, respectively.²⁵

Most of the hypotheses were confirmed in the results, but there were some surprising results as well. Language, irridentist potential, independent history, relative

Table 2

Log likelihood = -445.12391		Pseudo R2 = 0.3699			
Invote	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
rep	.0565759	.0104588	5.409	0.000 ***	.0360618 .07709
provelec	.4912995	.1270992	3.865	0.000 ***	.2420038 .7405953
enep	.419118	.0571172	7.338	0.000 ***	.3070867 .5311493
ideol	.0034196	.0087343	0.392	0.695	-.0137121 .0205513
left	.0018189	.0060294	0.302	0.763	-.0100073 .0136452
pop	-5.01e-06	.000019	-0.264	0.792	-.0000423 .0000322
misind	.0664968	.010759	6.181	0.000 ***	.0453938 .0875999
gdpratio	1.435217	.2290001	6.267	0.000 ***	.9860502 1.884384
unemdiff	-.0366022	.0190635	-1.920	0.055	-.0739938 .0007894
lang	.0261917	.002778	9.428	0.000 ***	.0207429 .0316406
langadj	-1.63672	.2787264	-5.872	0.000 ***	-2.183421 -1.090018
indep	1.160979	.2262523	5.131	0.000 ***	.7172018 1.604756
dist	1.023726	.1969388	5.198	0.000 ***	.6374447 1.410007
provauto	.1039376	.0597246	1.740	0.082	-.0132079 .2210831
dauto	.0352868	.1086277	0.325	0.745	-.1777786 .2483521
glob	.6716216	.4007733	1.676	0.094	-.114466 1.457709
_cons	-6.076787	.6642192	-9.149	0.000 ***	-7.379605 -4.77397

Obs. summary: 1473 left-censored observations at Invote<=1.1
166 uncensored observations

*** statistically significant from zero at the 99% level, two-tailed test
** statistically significant from zero at the 95% level, two-tailed test

Table 3

Log likelihood = -323.88438		Pseudo R2 = 0.5415			
Invote	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
lnlag	1.233533	.0663856	18.581	0.000 ***	1.103322 1.363743
rep	.0240575	.005824	4.131	0.000 ***	.0126342 .0354809
enep	.1523863	.0319291	4.773	0.000 ***	.0897597 .215013
ideol	-.0008335	.0049749	-0.168	0.867	-.0105914 .0089245
left	-.0020872	.0034576	-0.604	0.546	-.0088689 .0046946
pop	9.54e-07	.0000105	0.091	0.928	-.0000197 .0000216

misind		.0234211	.0060661	3.861	0.000	***	.011523	.0353193
gdpratio		.375978	.1292045	2.910	0.004	***	.1225529	.6294032
unemdiff		-.0220126	.0107305	-2.051	0.040	**	-.0430596	-.0009655
lang		.0066739	.0016056	4.157	0.000	***	.0035245	.0098232
langadj		-.3683841	.1517992	-2.427	0.015	**	-.666127	-.0706412
indep		.4224729	.1273188	3.318	0.001	***	.1727464	.6721994
dist		.6050596	.1097446	5.513	0.000	***	.3898036	.8203157
provauto		.0419965	.0333166	1.261	0.208		-.0233516	.1073445
dauto		.0034753	.0624201	0.056	0.956		-.1189572	.1259077
glob		-.0765532	.2302049	-0.333	0.740		-.5280832	.3749769
_cons		-2.772167	.3558504	-7.790	0.000	***	-3.470141	-2.074192
Obs. summary:		1473 left-censored observations at Invote<=1.1						
		166 uncensored observations						
***	statistically significant from zero at the 99% level, two-tailed test							
**	statistically significant from zero at the 95% level, two-tailed test							

affluence, provincial representation, and multi-partism all had the effects expected. The biggest surprises were the results on *DIST*, *UNEMDIFF*, and *MISIND*. It was certainly not expected that provinces with a high misery index would be more secessionist. This result turned out to be partially dependent on the Spanish observations, which are outliers on the *MISIND* variable. When the Spanish observations are dropped in the historical-regions dataset, both *UNEMDIFF* and *MISIND* switch signs. In the provinces dataset they remain statistically significant, though with smaller coefficients.²⁶ The interpretation is additionally complicated by the fact that in the regressions with linguistically distinct provinces and the Spanish observations dropped, both variables switch signs. The results reported here with the Spanish observations included seem to contradict the finding in fixed-effects time-series analysis that high unemployment correlates with low secessionist vote, but high relative unemployment correlates with high secessionist vote.²⁷ It seems likely

that the results on these variables, especially *MISIND*, are simply picking up the effect of some omitted variable.

The correlation found with *REP* also appears to be spurious, as the results reported in Tables 5 and 6 demonstrate. This variable should not have a direct effect on secessionism; its effect should be indirect and thus show up more clearly in the dataset limited to linguistically distinctive provinces. However, this does not turn out to be the case.

We see above that provincial autonomy, changes in provincial autonomy, and globalization do not seem to have important direct effects on secessionism. As Tables 5 and 6 show, the indirect effects are also weak at best.

It is natural now to ask which factors seem to be the most important in accounting for secessionist vote across regions. We can do this by using a regression equation to predict values of the dependent variable and see how the dependent variable changes based on changes in the independent variables.²⁸ In this exercise, the independent variables *GDP**RATIO* and *ENEP* are fixed at their 80th percentile values (that is, the values at which 80 per cent of the observations take on smaller values) then moved to their 20th percentile values, while the dummy variables are switched between 0 and 1.²⁹ *LANG*'s 80th percentile value (0) is actually lower than its mean (6.3), indicating a highly skewed distribution. This variable is

therefore fixed at 50, then moved to its mean of 6.3 to establish its substantive influence.

Table 4 presents different values for variables of interest and examines how secessionist vote changes when the independent variables are changed. For all of these counterfactual predictions, *REP*, *IDEOL*, *LEFT*, *POP*, *MISIND*, and *UNEMDIFF* have been set to their sample means, while *PROVAUTO* has been set to 3, *DAUTO* to 0, and *GLOB* to its 1999 value, 1.45. The regression equation used is that presented in Table 2.

Table 4

PROVELEC	ENEP	GDP	RATIO	LANG	LANGADJ	INDEP	DIST	LN	VOTE	VOTE
0	3.7	1.10	50	0	0	0	0	0.811	-0.7	
0	3.7	1.10	50	0	0	0	1	1.835	3.3	
0	3.7	1.10	50	0	1	0	0	1.972	4.2	
1	3.7	1.10	50	0	0	0	0	1.303	0.7	
1	3.7	1.10	50	0	1	1	1	3.487	29.7	
1	3.7	1.10	50	1	1	1	1	1.851	3.4	
1	3.7	0.82	50	0	1	1	1	3.081	18.8	
1	2.0	1.10	50	0	1	1	1	2.766	12.9	
1	3.7	1.10	6.3	0	1	1	1	2.343	7.4	

As the above figures demonstrate, the perturbation of *LANGADJ* has quite substantial effects compared to some of the other variables. When *PROVELEC*, *LANGADJ*, *INDEP*, and *DIST* are set to their most “favorable” values for secessionism, predicted secessionist vote is about 29.7%. But when the province is given irridentist potential, secessionist vote drops all the way to 3.4%. By contrast, perturbing *GDP* *RATIO* from its 80th to its 20th percentile value reduces secessionist vote only to 18.8%, while doing the same with *ENEP*

reduces secessionist vote to 12.9%. Even perturbing minority language speakers from 50% of the population to 6.3% reduces secessionist vote only to 7.4% under these conditions.³⁰

So much for the results on the full dataset. Tables 5 and 6 report the results for the regressions on the dataset limited to linguistically distinctive provinces.

Table 5

Log likelihood = -223.08865		Pseudo R2 = 0.2967			
Invote	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
rep	.0078244	.0145928	0.536	0.592	-.0209062 .036555
provelec	.1185222	.1643366	0.721	0.471	-.2050273 .4420718
enep	.3888745	.0840243	4.628	0.000 ***	.2234456 .5543033
ideol	.0069152	.0115006	0.601	0.548	-.0157273 .0295578
left	.009283	.0095518	0.972	0.332	-.0095228 .0280887
pop	.0003399	.0000696	4.884	0.000 ***	.0002029 .0004769
misind	.0391673	.0138787	2.822	0.005 ***	.0118425 .066492
gdpratio	.9538319	.2466951	3.866	0.000 ***	.4681332 1.43953
unemdiff	-.0153884	.0214426	-0.718	0.474	-.057605 .0268281
lang	.0045338	.0040492	1.120	0.264	-.0034384 .012506
langadj	-1.228917	.2636502	-4.661	0.000 ***	-1.747998 -.7098369
indep	-.7474545	.3998322	-1.869	0.063	-1.534653 .0397439
dist	.953893	.2714556	3.514	0.001 ***	.4194452 1.488341
provauto	.0930332	.0782334	1.189	0.235	-.0609944 .2470607
dauto	.0803307	.1242858	0.646	0.519	-.1643658 .3250273
glob	.9109652	.5096308	1.788	0.075	-.092407 1.914337
_cons	-4.095107	.7905503	-5.180	0.000 ***	-5.651559 -2.538654

Obs. summary: 170 left-censored observations at Invote<=1.1
115 uncensored observations

*** statistically significant from zero at the 99% level, two-tailed test
** statistically significant from zero at the 95% level, two-tailed test

Table 6

Log likelihood = -140.76068		Pseudo R2 = 0.5563			
Invote	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
lnlag	1.035444	.0635533	16.293	0.000 ***	.910319 1.160569
rep	-.0003242	.0079819	-0.041	0.968	-.0160392 .0153908
enep	.1113697	.0460786	2.417	0.016 **	.0206492 .2020901
ideol	-.0003546	.0062081	-0.057	0.954	-.0125771 .011868
left	.001868	.0051881	0.360	0.719	-.0083465 .0120824
pop	.0001342	.0000382	3.509	0.001 ***	.0000589 .0002094
misind	.0038557	.007595	0.508	0.612	-.0110975 .0188088
gdpratio	.3137139	.1347261	2.329	0.021 **	.0484622 .5789655
unemdiff	-.0090602	.0116966	-0.775	0.439	-.0320887 .0139683
lang	.0013757	.0021944	0.627	0.531	-.0029446 .0056961
langadj	-.4078523	.1428907	-2.854	0.005 ***	-.6891787 -.126526
indep	-.5362539	.2235697	-2.399	0.017 **	-.9764228 -.096085
dist	.5230996	.147868	3.538	0.000 ***	.2319739 .8142253
provauto	-.0118041	.0424829	-0.278	0.781	-.0954455 .0718372
dauto	.0539118	.066087	0.816	0.415	-.0762017 .1840253

glob		-.0117017	.2826182	-0.041	0.967	-.5681267	.5447232
_cons		-1.426421	.4252138	-3.355	0.001 ***	-2.263591	-.5892512
Obs. summary:		170 left-censored observations at lnvote<=1.1					
		115 uncensored observations					
***		statistically significant from zero at the 99% level, two-tailed test					
**		statistically significant from zero at the 95% level, two-tailed test					

These results are largely as expected: variables supposed to have a strong conditional effect, such as irridentist potential (*LANGADJ*), population, and (perhaps) multi-partism indeed tend to have stronger effects in the sample limited to linguistically distinctive provinces than in the full sample.³¹ By contrast, variables with a clear direct effect, such as language and independent history, have lower coefficients (the coefficient on *INDEP* is actually negative). The regression results reported here actually seem to exaggerate the extent to which language has a smaller effect among provinces that are already linguistically distinctive. The results presented in Tables 5 and 6 would suggest that a province need only have a few speakers of a minority language in order to be fertile ground for secessionism: additional numbers of speakers would not matter. However, the results for the regressions on the historical-regions dataset and with the *LNVOTE2* variable both show an effect of language even when the datasets are limited, thus suggesting that linguistically distinctive provinces with greater numbers of minority language speakers have somewhat more successful secessionist parties. Still, the difference in results between the two datasets is striking, and we must conclude that the effects of language on secessionism are nonlinear: the

greatest effect comes with the first few speakers, and the effect declines thereafter. The coefficients on *GDPRATIO* were modestly lower in the limited sample, suggesting that its effects are both direct and indirect. Relative affluence can be a substitute for linguistic difference in promoting secessionism, but it can also work with linguistic difference.

Conclusion

What are the main “risk factors” for secessionist electoral success? A region-specific minority language (number of speakers does not appear to matter as much as simply having a living regional language), non-contiguity of region with independent country in which regional language is an official language (lack of irridentist potential), history of independence, non-contiguity of region with rest of country, relative affluence of region, absolute size of region in terms of population, and multi-party political system are shown consistently to have important effects on the level of observed secessionism. Some of these variables, particularly population, but also multi-party political system and, of course, lack of irridentist potential, show a highly conditional effect: they do not in themselves stimulate secessionism, but they allow for it when the region is already linguistically distinctive. The result on non-contiguity of region with the rest of the country was a surprise. Non-contiguity has certainly aided secessionist rebellions and was crucial in Bangladesh’s independence

bid. But in democracies where secession would not be militarily crushed, it seems strange that non-contiguity would have such importance. Perhaps this variable is actually capturing cultural and economic differences in remote areas that are not captured by variables like language, independent history, and ideological difference. When a region is separated from the rest of the country by sea, its psychological bond with the country may be less, its cultural development may be different, and its trade might rely more on international than domestic markets. Geographical separateness might encourage a sense of separateness along many other dimensions as well, providing fertile ground for secessionism.³²

The following factors were not found to be consistently important in explaining secessionism across regions: relative size of province within country, provincial autonomy, change in provincial autonomy, provincial misery index, relative provincial unemployment, globalization, absolute ideological difference, and relative left-wing orientation. In a regression format that emphasizes the time-series variation and limits the sample to provinces with well-established secessionist parties, some of these factors do seem to be important, but unfortunately there is insufficient space here to present that research.

Is it possible on the basis of this research to hazard some predictions about where secessionism should be appearing next – or

where it should be drying up? One way to attempt to use the results in this way is to look at observations which the model does not predict well. If a province repeatedly displays a vote share below that predicted, we may conclude one of two things: either there is an omitted variable that would explain the discrepancy if it had been included in the regression, or the province is “out of equilibrium,” that is, it really should have more secessionism than currently exists there – a political entrepreneur could presumably exploit the issue to effect. If a province repeatedly displays a secessionist vote share above that predicted, then it is most likely that an omitted variable accounts for the discrepancy, but it is possible here as well that the province is out of equilibrium and ripe for a downturn in secessionism.

For the purposes of these calculations, the results for regressions without a lagged vote variable should be used, since the lagged vote variable takes into account a province’s “starting points” in terms of electoral secessionism. To make the calculations, I have used both the regressions presented in Tables 2 and 5. The tables of the 15 most overpredicted and underpredicted observations in both datasets are presented in the Appendix (four tables in all). There were some consistent findings: secessionism was underpredicted in Scotland and Wales, and overpredicted in Sardinia, Flanders, and Wallonia. Secessionism also tended to be underpredicted in Alaska in the late 1980s and throughout the 1990s but overpredicted there in

the early 1980s, when Alaska's *GDP/RATIO* skyrocketed briefly due to the boom in oil prices. This time trend seems to indicate that Alaskan secessionism is due for decline or even disappearance.

With regard to Scotland and Wales, it is likely that some omitted variable is at work. Popular nationalism in Scotland and Wales developed early, perhaps earlier than elsewhere in Europe, and has been remarkably resilient, even when they have gone centuries without their own popularly elected regional institutions. There is something about Scottish and Welsh nationalism that the variables included in this paper's models do not explain.

With regard to Sardinia, secessionism is so overpredicted there because at least 60% of the population of the island speak Sard, yet the language is excluded from Sardinian public life and is often denigrated as a dialect, even though linguists consider it a distinct language that developed directly from Latin. A long-running moderate secessionist party does exist in Sardinia, the Party of Sardinian Action, sometimes joined by the radical secessionist Sardinian Independentist Party, but both parties have done rather poorly in elections. The problem seems to be that a Sardinian national identity based on the Sard language has not yet developed. Were such an identity to develop, secessionism should increase among the island's voters.

In Flanders and Wallonia as well, an increase in secessionism seems highly possible. Secessionism in both areas is highly conditional. If Flanders is able to receive full fiscal autonomy and a greater status for the Dutch language in Brussels, it will likely continue to reject secessionism, but if Flanders receives fiscal federalism, Wallonia might pull out. Thus, the issue could well reach a tipping point beyond which both communities view secession as the only option. Indeed, polls suggest that Flemish electoral support for the Flemish Bloc and the Volksunie understates support for secession, while secessionist Walloons are without any party that supports their views. An August 1996 poll of 1,000 Belgians published in the Walloon newspaper *La Nouvelle Gazette* gave the following figures for support for various constitutional options: “Separatism” was favored by 34.5% of Flemish interviewees, 22.4% of Walloon interviewees, and 24.5% of Brussels interviewees, while “Federalism” was favored by 59.6% of Flemings, 54.7% of Walloons, and 47.1% of Brussels residents. “Unitarism” was favored by just 5.9% of Flemings, 22.9% of Walloons, and 28.4% of Brussels residents.³³

If recent trends are any indication, secessionism is here to stay in Western democracies. This paper has presented new research uncovering the cross-sectional determinants of secessionism. This research should not only help us to understand secessionism better, but should also open the door to wider comparative study of regional

politics and how central governments respond to regional demands through offers of autonomy, fiscal concessions, and so on.

Appendix

15 province-years for which secessionism was most underpredicted, full-sample regression (not including provinces with zero secessionist vote but which were underpredicted due to the censoring point):

Province	Year	Fitted Value - <i>LNVOTE</i>
Alaska	1992	-3.775306
Wales	1999	-3.623899
Wales	1997	-3.471108
Alaska	1996	-3.239029
Wales	1992	-3.217018
Liguria	1992	-3.191298
Wales	1983	-3.189941
Alaska	1998	-2.963307
Alaska	1986	-2.934108
Scotland	1997	-2.933855
Scotland	1992	-2.806502
Wales	1987	-2.795542
Lombardy	1992	-2.790230
Alberta	1982	-2.751992
Piedmont	1992	-2.706732

15 province-years for which secessionism was most overpredicted, full-sample regression:

Province	Year	Fitted Value - <i>LNVOTE</i>
Sardinia	1984	1.791414
Sardinia	1989	1.631096
Sardinia	1996	1.530264
Sardinia	1994	1.409127
Flanders	1995	1.297643
Aaland	1995	1.090727
Sardinia	1987	1.032320
Flanders	1985	1.002655
Sardinia	1983	.9766327
Sardinia	1992	.9180005
Wallonia	1995	.8440588
Flanders	1995	.8063438
Sardinia	1994	.7320359
Quebec	1988	.6443646
Geneva	1997	.6260751

Note: two values for a single province-year indicate that one is a national election, the other provincial.

15 province-years for which secessionism was most underpredicted, linguistically distinctive-sample regression (not including provinces with zero secessionist vote but which were underpredicted due to the censoring point):

Province	Year	Fitted Value - <i>LNVOTE</i>
Alaska	1992	-2.272018
Wales	1999	-2.043149
Scotland	1992	-2.025918
Friuli-Venezia Giulia	1993	-1.992588
Scotland	1997	-1.971340
Scotland	1983	-1.842580
Faroe Islands	1998	-1.834072
Wales	1983	-1.764719
Scotland	1999	-1.716589
Wales	1997	-1.700774
Friuli-Venezia Giulia	1992	-1.643360
Scotland	1987	-1.620718
Wales	1992	-1.619148
Alaska	1986	-1.613801
Alaska	1996	-1.566441

15 province-years for which secessionist was most overpredicted, linguistically distinctive-sample regression:

Province	Year	Fitted Value - <i>LNVOTE</i>
Sardinia	1996	1.473435
Sardinia	1994	1.323913
Quebec	1988	1.300570
Rhone-Alpes	1992	1.158465
Sardinia	1984	1.078210
Sardinia	1989	1.070866
Sardinia	1987	.8690574
Provence-Alpes-Cote d'Azur	1992	.8059543
Sardinia	1992	.7984762
Hokkaido	1995	.7452874
Quebec	1984	.7445529
Alaska	1982	.6409614
Rhone-Alpes	1986	.6370764
Alaska	1980	.4485743
Provence-Alpes-Cote d'Azur	1986	.3588510

Additional Regression Results

Full provinces dataset, no lagged vote variable, *LNVOTE2* dependent variable:

Log likelihood = -449.68426		Pseudo R2 = 0.3878			
lnvote2	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
rep	.0591415	.0109045	5.424	0.000 ***	.0377531 .0805299
provelec	.5428699	.1332506	4.074	0.000 ***	.2815086 .8042311
enep	.454569	.0596301	7.623	0.000 ***	.337609 .571529
ideol	.0010354	.0092299	0.112	0.911	-.0170683 .019139
left	-.0011719	.0063873	-0.183	0.854	-.0137 .0113563
pop	-4.14e-06	.0000197	-0.210	0.834	-.0000428 .0000345
misind	.0793269	.0111899	7.089	0.000 ***	.0573789 .101275
gdpratio	1.766766	.2381566	7.419	0.000 ***	1.299639 2.233892
unemdiff	-.0368293	.0198755	-1.853	0.064	-.0758137 .0021551
lang	.0318592	.0028976	10.995	0.000 ***	.0261757 .0375427
langadj	-2.044789	.2925436	-6.990	0.000 ***	-2.618592 -1.470987
indep	1.260763	.2374689	5.309	0.000 ***	.7949852 1.726541
dist	1.014881	.2063392	4.919	0.000 ***	.6101616 1.4196
provauto	.1068663	.0626576	1.706	0.088	-.016032 .2297646
dauto	.0199722	.1141227	0.175	0.861	-.2038711 .2438155
glob	.7510205	.4210106	1.784	0.075	-.074761 1.576802
_cons	-6.903877	.6939948	-9.948	0.000 ***	-8.265097 -5.542657

Obs. summary: 1473 left-censored observations at lnvote2<=1.1
166 uncensored observations

*** statistically significant from zero at the 99% level, two-tailed test
** statistically significant from zero at the 95% level, two-tailed test

Full provinces dataset, lagged vote variable, *LNVOTE2* dependent variable:

Log likelihood = -339.64671		Pseudo R2 = 0.5376			
lnvote2	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
lnlag2	1.29142	.0747236	17.283	0.000 ***	1.144855 1.437984
rep	.026447	.0065682	4.027	0.000 ***	.0135639 .0393301
enep	.1852093	.0361569	5.122	0.000 ***	.1142901 .2561284
ideol	-.0039784	.0057394	-0.693	0.488	-.0152358 .007279
left	-.0058791	.0040395	-1.455	0.146	-.0138024 .0020442
pop	2.08e-06	.0000117	0.177	0.859	-.0000209 .0000251
misind	.0371246	.0067921	5.466	0.000 ***	.0238023 .0504469
gdpratio	.6853872	.1455256	4.710	0.000 ***	.3999494 .9708249
unemdiff	-.0226527	.012087	-1.874	0.061	-.0463604 .001055
lang	.0120817	.0018148	6.657	0.000 ***	.0085221 .0156413
langadj	-.7213526	.1764145	-4.089	0.000 ***	-1.067377 -.3753285
indep	.5202089	.145449	3.577	0.000 ***	.2349213 .8054965
dist	.613055	.1248649	4.910	0.000 ***	.3681417 .8579682
provauto	.0462331	.0379191	1.219	0.223	-.0281424 .1206085
dauto	-.0130792	.0707522	-0.185	0.853	-.1518544 .1256961
glob	-.0168325	.2627968	-0.064	0.949	-.5322892 .4986242
_cons	-3.659262	.4041646	-9.054	0.000 ***	-4.452002 -2.866523

Obs. summary: 1473 left-censored observations at lnvote2<=1.1
166 uncensored observations

*** statistically significant from zero at the 99% level, two-tailed test
** statistically significant from zero at the 95% level, two-tailed test

**“Linguistically distinctive” provinces dataset, no lagged vote variable,
LNVOTE2 dependent variable:**

Log likelihood = -224.40623 Pseudo R2 = 0.3406

lnvote2	Coef.	Std. Err.	t	P> t		[95% Conf. Interval]
rep	.0017356	.0151597	0.114	0.909		-.0281111 .0315824
provelec	.1530143	.1713393	0.893	0.373		-.1843222 .4903508
enep	.4920564	.0880666	5.587	0.000	***	.3186689 .6654439
ideol	.0038637	.0120432	0.321	0.749		-.0198471 .0275746
left	.0043168	.0100357	0.430	0.667		-.0154417 .0240753
pop	.0004288	.000073	5.870	0.000	***	.000285 .0005726
misind	.0558974	.0143779	3.888	0.000	***	.0275898 .084205
gdpratio	1.246962	.2559789	4.871	0.000	***	.7429852 1.750939
unemdiff	-.0184549	.0222948	-0.828	0.409		-.0623494 .0254396
lang	.0073548	.0042536	1.729	0.085		-.0010198 .0157295
langadj	-1.466823	.2794057	-5.250	0.000	***	-2.016923 -.9167227
indep	-.6035101	.4181285	-1.443	0.150		-1.426731 .2197104
dist	.9964263	.2850996	3.495	0.001	***	.435116 1.557737
provauto	.1109657	.0821359	1.351	0.178		-.0507452 .2726766
dauto	.0590351	.128921	0.458	0.647		-.1947874 .3128575
glob	1.007136	.5339275	1.886	0.060		-.0440719 2.058344
_cons	-5.283755	.8251829	-6.403	0.000	***	-6.908393 -3.659117

Obs. summary: 170 left-censored observations at lnvote2<=1.1
115 uncensored observations

*** statistically significant from zero at the 99% level, two-tailed test
** statistically significant from zero at the 95% level, two-tailed test

**“Linguistically distinctive” provinces dataset, lagged vote variable,
LNVOTE2 dependent
variable:**

Log likelihood = -156.25695 Pseudo R2 = 0.5409

lnvote2	Coef.	Std. Err.	t	P> t		[95% Conf. Interval]
lnlag2	1.051641	.0764129	13.763	0.000	***	.9011977 1.202085
rep	-.0087874	.0096727	-0.908	0.364		-.0278313 .0102565
enep	.2302509	.0564987	4.075	0.000	***	.119015 .3414867
ideol	-.0041609	.0076073	-0.547	0.585		-.0191382 .0108165
left	-.0033362	.0064063	-0.521	0.603		-.0159491 .0092768
pop	.0002491	.0000471	5.287	0.000	***	.0001563 .0003418
misind	.0219563	.0090976	2.413	0.016	**	.0040448 .0398678
gdpratio	.6299241	.1614012	3.903	0.000	***	.3121539 .9476943
unemdiff	-.0138643	.0141744	-0.978	0.329		-.0417712 .0140425
lang	.0041521	.0026867	1.545	0.123		-.0011375 .0094417
langadj	-.6423468	.1787826	-3.593	0.000	***	-.9943378 -.2903557
indep	-.5005708	.2779466	-1.801	0.073		-1.047798 .0466565
dist	.6281672	.1821674	3.448	0.001	***	.269512 .9868224
provauto	.0031904	.0526393	0.061	0.952		-.1004471 .1068278
dauto	.0466188	.0795525	0.586	0.558		-.1100059 .2032436
glob	.0797506	.3466791	0.230	0.818		-.6027987 .7622999
_cons	-2.813366	.5185188	-5.426	0.000	***	-3.834237 -1.792494

Obs. summary: 170 left-censored observations at lnvote2<=1.1
115 uncensored observations

*** statistically significant from zero at the 99% level, two-tailed test
** statistically significant from zero at the 95% level, two-tailed test

Full historical-regions dataset, no lagged vote variable, *LNVOTE*
dependent variable:³⁴

Log likelihood = -300.57466		Pseudo R2 = 0.2622			
lnvote	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
rep	.0185028	.0082711	2.237	0.026 **	.0022458 .0347598
regelec	.2961303	.1520289	1.948	0.052	-.0026858 .5949464
enep	.2574386	.0751975	3.423	0.001 ***	.1096362 .405241
ideol	.0064626	.0103693	0.623	0.533	-.0139186 .0268438
left	.0082822	.0076318	1.085	0.278	-.0067183 .0232828
pop	-.000014	9.72e-06	-1.440	0.151	-.0000331 5.11e-06
misind	.038208	.0119235	3.204	0.001 ***	.014772 .061644
gdpratio	1.246823	.2287979	5.449	0.000 ***	.7971155 1.69653
unemdiff	.002963	.019115	0.155	0.877	-.0346079 .040534
lang	.0234919	.0028646	8.201	0.000 ***	.0178614 .0291224
langadj	-1.163653	.2547683	-4.567	0.000 ***	-1.664406 -.6629005
indep	.5170469	.2543553	2.033	0.043 **	.0171059 1.016988
dist	.1697742	.1939972	0.875	0.382	-.2115315 .5510799
regauto	.3065602	.1445639	2.121	0.035 **	.0224167 .5907037
dregauto	.4610669	.2157318	2.137	0.033 **	.0370413 .8850925
glob	.6534004	.4507837	1.449	0.148	-.2326249 1.539426
_cons	-4.291277	.7305389	-5.874	0.000 ***	-5.727167 -2.855386

Obs. summary: 311 left-censored observations at lnvote<=1.1
133 uncensored observations

*** statistically significant from zero at the 99% level, two-tailed test
** statistically significant from zero at the 95% level, two-tailed test

Full historical-regions dataset, lagged vote variable, *LNVOTE*
dependent variable:

Log likelihood = -182.21988		Pseudo R2 = 0.5527			
lnvote	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
lnlag	1.154199	.0590677	19.540	0.000 ***	1.0381 1.270298
rep	.0117247	.0041235	2.843	0.005 ***	.0036199 .0198295
enep	.01517	.0375609	0.404	0.687	-.0586567 .0889967
ideol	.0026098	.0049828	0.524	0.601	-.007184 .0124036
left	-.0043236	.0037973	-1.139	0.256	-.0117873 .0031401
pop	-6.66e-06	4.67e-06	-1.428	0.154	-.0000158 2.51e-06
misind	.0140953	.0058226	2.421	0.016 **	.0026508 .0255398
gdpratio	.1936528	.1149467	1.685	0.093	-.0322774 .419583
unemdiff	-.0011379	.0093159	-0.122	0.903	-.0194485 .0171726
lang	.0053659	.0014476	3.707	0.000 ***	.0025206 .0082111
langadj	-.1864013	.1196059	-1.558	0.120	-.4214893 .0486867
indep	.2083473	.125717	1.657	0.098	-.0387522 .4554469
dist	.245572	.0969077	2.534	0.012 **	.0550978 .4360463
regauto	.0905587	.0690607	1.311	0.190	-.0451816 .226299
dregauto	.2332321	.1020884	2.285	0.023 **	.0325751 .4338891
glob	.0031943	.2235676	0.014	0.989	-.4362327 .4426213
_cons	-1.656161	.3494523	-4.739	0.000 ***	-2.343017 -.9693047

Obs. summary: 311 left-censored observations at lnvote<=1.1
133 uncensored observations

*** statistically significant from zero at the 99% level, two-tailed test
** statistically significant from zero at the 95% level, two-tailed test

“Linguistically distinctive” historical-regions dataset, no lagged vote variable, *LNVOTE* dependent variable:

Log likelihood = -218.88244		Pseudo R2 = 0.2664			
lnvote	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
rep	.0026658	.0133542	0.200	0.842	-.0236407 .0289722
regelec	.2489423	.1759224	1.415	0.158	-.0976068 .5954914
enep	.3083156	.0808317	3.814	0.000 ***	.1490855 .4675458
ideol	.0027826	.0127117	0.219	0.827	-.0222581 .0278232
left	.0230438	.0102475	2.249	0.025 **	.0028572 .0432303
pop	.0001234	.0000361	3.413	0.001 ***	.0000522 .0001946
misind	.03833	.0146454	2.617	0.009 ***	.00948 .06718
unemdiff	.0120265	.0215128	0.559	0.577	-.0303515 .0544045
gdpratio	1.151113	.2639528	4.361	0.000 ***	.631153 1.671073
lang	.0142126	.0040946	3.471	0.001 ***	.0061467 .0222786
langadj	-.9654095	.2594535	-3.721	0.000 ***	-1.476506 -.4543127
indep	-.5805856	.391783	-1.482	0.140	-1.352358 .1911868
dist	.3285922	.2623884	1.252	0.212	-.1882861 .8454705
regauto	.1645082	.1637663	1.005	0.316	-.1580946 .4871109
dregauto	.4323764	.2670378	1.619	0.107	-.0936608 .9584136
glob	.9521521	.5240955	1.817	0.071	-.0802625 1.984567
_cons	-3.943721	.8605684	-4.583	0.000 ***	-5.638952 -2.248489

Obs. summary: 145 left-censored observations at lnvote<=1.1
111 uncensored observations

*** statistically significant from zero at the 99% level, two-tailed test
** statistically significant from zero at the 95% level, two-tailed test

“Linguistically distinctive” historical-regions dataset, lagged vote variable, *LNVOTE* dependent variable:

Log likelihood = -124.2612		Pseudo R2 = 0.5835			
lnvote	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
lnlag	1.092761	.061705	17.709	0.000 ***	.971208 1.214313
rep	.0002007	.0066366	0.030	0.976	-.0128727 .013274
enep	.0297398	.041396	0.718	0.473	-.051806 .1112856
ideol	-.0035351	.0062083	-0.569	0.570	-.0157648 .0086947
left	.0056417	.0050276	1.122	0.263	-.0042622 .0155455
pop	.0000623	.000018	3.452	0.001 ***	.0000267 .0000978
misind	.0076553	.0072948	1.049	0.295	-.0067148 .0220254
gdpratio	.2749977	.1309924	2.099	0.037 **	.0169562 .5330393
unemdiff	-.0041951	.0106989	-0.392	0.695	-.0252708 .0168805
lang	.0041266	.0020612	2.002	0.046 **	.0000662 .008187
langadj	-.1444192	.1327817	-1.088	0.278	-.4059856 .1171473
indep	-.3584313	.199119	-1.800	0.073	-.7506753 .0338126
dist	.3752703	.1331387	2.819	0.005 ***	.1130007 .6375399
regauto	-.0276807	.081664	-0.339	0.735	-.1885504 .1331891
dregauto	.2781325	.1233683	2.254	0.025 **	.0351095 .5211555
glob	-.0167238	.2653139	-0.063	0.950	-.539365 .5059173
_cons	-1.199046	.4256935	-2.817	0.005 ***	-2.037619 -.3604737

Obs. summary: 145 left-censored observations at lnvote<=1.1
111 uncensored observations

*** statistically significant from zero at the 99% level, two-tailed test
** statistically significant from zero at the 95% level, two-tailed test

Full historical-regions dataset, no lagged vote variable, *LNVOTE2*
dependent variable:

Log likelihood = -303.69595 Pseudo R2 = 0.2940

lnvote2	Coef.	Std. Err.	t	P> t		[95% Conf. Interval]
rep	.0126823	.0089837	1.412	0.159		-.0049754 .0303401
regelec	.3947629	.165777	2.381	0.018	**	.0689246 .7206011
enep	.3264131	.0817918	3.991	0.000	***	.1656495 .4871768
ideol	.000624	.011411	0.055	0.956		-.0218045 .0230525
left	.0084248	.008474	0.994	0.321		-.008231 .0250806
pop	-.0000117	.0000104	-1.129	0.259		-.0000322 8.70e-06
misind	.0509902	.0129286	3.944	0.000	***	.0255787 .0764017
gdpratio	1.765286	.2517164	7.013	0.000	***	1.270532 2.26004
unemdiff	.011329	.0207803	0.545	0.586		-.0295151 .0521732
lang	.0309905	.0031756	9.759	0.000	***	.0247489 .0372322
langadj	-1.547075	.2797666	-5.530	0.000	***	-2.096962 -.9971873
indep	.6940198	.2781292	2.495	0.013	**	.1473506 1.240689
dist	-.0110792	.2132034	-0.052	0.959		-.4301351 .4079768
regauto	.2743613	.1576039	1.741	0.082		-.0354127 .5841352
dregauto	.4173322	.2337593	1.785	0.075		-.0421269 .8767913
glob	.7066518	.4914552	1.438	0.151		-.2593143 1.672618
_cons	-5.271831	.7969485	-6.615	0.000	***	-6.838251 -3.705411

Obs. summary: 313 left-censored observations at lnvote2<=1.1
131 uncensored observations

*** statistically significant from zero at the 99% level, two-tailed test
** statistically significant from zero at the 95% level, two-tailed test

Full historical-regions dataset, lagged vote variable, *LNVOTE2*
dependent variable:

Log likelihood = -181.12443 Pseudo R2 = 0.5789

lnvote2	Coef.	Std. Err.	t	P> t		[95% Conf. Interval]
lnlag2	1.085587	.0511611	21.219	0.000	***	.9850283 1.186145
rep	.0113491	.004133	2.746	0.006	***	.0032255 .0194727
enep	.0401111	.03768	1.065	0.288		-.0339497 .1141719
ideol	.0014284	.005031	0.284	0.777		-.0084603 .011317
left	-.0021062	.0037782	-0.557	0.578		-.0095323 .0053199
pop	-5.94e-06	4.57e-06	-1.301	0.194		-.0000149 3.04e-06
misind	.0130909	.0058332	2.244	0.025	**	.0016256 .0245562
gdpratio	.2163691	.122381	1.768	0.078		-.0241735 .4569118
unemdiff	.000976	.0093458	0.104	0.917		-.0173934 .0193453
lang	.0064681	.0015244	4.243	0.000	***	.0034719 .0094644
langadj	-.2086948	.1230886	-1.695	0.091		-.4506282 .0332385
indep	.2377244	.1256365	1.892	0.059		-.0092169 .4846657
dist	.2630422	.0986906	2.665	0.008	***	.0690636 .4570209
regauto	.0854868	.068996	1.239	0.216		-.0501263 .2210999
dregauto	.2515699	.1013786	2.481	0.013	**	.0523081 .4508318
glob	.0071051	.223012	0.032	0.975		-.43123 .4454402
_cons	-1.666785	.351779	-4.738	0.000	***	-2.358215 -.9753561

Obs. summary: 313 left-censored observations at lnvote2<=1.1
131 uncensored observations

*** statistically significant from zero at the 99% level, two-tailed test
** statistically significant from zero at the 95% level, two-tailed test

“Linguistically distinct” historical-regions dataset, no lagged vote variable, *LNVOTE2* dependent variable:

Log likelihood = -225.29399				Pseudo R2 =		0.2956	
lnvote2	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]		
rep	-.0043617	.0144135	-0.303	0.762	-.0327549	.0240315	
regelec	.3091931	.1899803	1.628	0.105	-.0650487	.6834349	
enep	.3783173	.0872889	4.334	0.000 ***	.2063671	.5502675	
ideol	-.0000795	.0137823	-0.006	0.995	-.0272292	.0270703	
left	.0221467	.0111354	1.989	0.048 **	.000211	.0440823	
pop	.0001379	.0000391	3.524	0.001 ***	.0000608	.0002149	
misind	.0550758	.0157577	3.495	0.001 ***	.0240347	.0861169	
unemdiff	.0185356	.0231604	0.800	0.424	-.027088	.0641592	
gdpratio	1.541057	.2848955	5.409	0.000 ***	.9798419	2.102272	
lang	.0197162	.0044248	4.456	0.000 ***	.0109999	.0284326	
langadj	-1.270749	.2829001	-4.492	0.000 ***	-1.828033	-.7134644	
indep	-.526143	.4270353	-1.232	0.219	-1.367359	.3150729	
dist	.1719872	.2828684	0.608	0.544	-.3852346	.729209	
regauto	.1665156	.1773828	0.939	0.349	-.1829103	.5159414	
dregauto	.3812719	.2880395	1.324	0.187	-.1861365	.9486803	
glob	1.002775	.5678256	1.766	0.079	-.1157829	2.121334	
_cons	-4.889782	.9283821	-5.267	0.000 ***	-6.718599	-3.060964	

Obs. summary: 145 left-censored observations at lnvote2<=1.1
111 uncensored observations

*** statistically significant from zero at the 99% level, two-tailed test
** statistically significant from zero at the 95% level, two-tailed test

“Linguistically distinct” historical-regions dataset, lagged vote variable, *LNVOTE2* dependent variable:

Log likelihood = -122.08161				Pseudo R2 =		0.6183	
lnvote2	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]		
lnlag2	1.039746	.0528282	19.682	0.000 ***	.9356804	1.143813	
rep	-.0001521	.0065141	-0.023	0.981	-.0129842	.0126799	
enep	.0514957	.0410281	1.255	0.211	-.0293254	.1323168	
ideol	-.0046992	.0061424	-0.765	0.445	-.016799	.0074006	
left	.007258	.0049792	1.458	0.146	-.0025504	.0170665	
pop	.0000631	.0000177	3.571	0.000 ***	.0000283	.0000979	
misind	.0061972	.0072512	0.855	0.394	-.0080869	.0204812	
gdpratio	.2745696	.1329371	2.065	0.040 **	.0126971	.5364422	
unemdiff	-.0033628	.0105485	-0.319	0.750	-.0241422	.0174166	
lang	.0050773	.0020416	2.487	0.014 **	.0010555	.0090991	
langadj	-.1649379	.1334418	-1.236	0.218	-.4278046	.0979288	
indep	-.3318357	.1938053	-1.712	0.088	-.7136123	.0499409	
dist	.4106361	.1332516	3.082	0.002 ***	.148144	.6731281	
regauto	-.0343355	.0801195	-0.429	0.669	-.1921627	.1234917	
dregauto	.2950512	.1211274	2.436	0.016 **	.0564426	.5336597	
glob	-.0672327	.261181	-0.257	0.797	-.5817326	.4472672	
_cons	-1.124063	.4225407	-2.660	0.008 ***	-1.956425	-.2917011	

Obs. summary: 145 left-censored observations at lnvote2<=1.1
111 uncensored observations

*** statistically significant from zero at the 99% level, two-tailed test
** statistically significant from zero at the 95% level, two-tailed test

Definition of Historical Regions

HISTORICAL REGION	CORRESPONDING ADMINISTRATIVE UNIT(S)	HISTORICAL REGION	CORRESPONDING ADMINISTRATIVE UNIT(S)
Northern Territory	Northern Territory (Australia)	Padania	Liguria, Piemonte, Lombardia, Veneto, Friuli-Venezia Giulia (minus Trieste), Trentino-Alto Adige (minus Bolzano) (Italy)
Queensland	Queensland (Australia)	Sardinia	Sardegna (Italy)
Tasmania	Tasmania (Australia)	Sicily	Sicilia (Italy)
Western Australia	Western Australia (Australia)	Southern Italy	Campania, Puglia, Basilicata, Calabria (Italy)
Flanders	Flanders (Belgium)	Valle d'Aoste	Valle d'Aosta (Italy)
Wallonia	Wallonia (Belgium)	Hokkaido	Hokkaido (Japan)
Arctic	Yukon Terr., Northwest Terr., Nunavut (Canada)	Kyushu	Prefectures of Kyushu Island (Japan)
Midwest Canada	Saskatchewan, Manitoba (Canada)	Okinawa	Ryukyu (Japan)
Arcadia	New Brunswick (Canada)	Shikoku	Prefectures of Shikoku Island (Japan)
Newfoundland	Newfoundland (Canada)	West Friesland	Friesland (Netherlands)
Atlantic Provinces	Nova Scotia, Prince Edward Island (Canada)	West Lappland	Finnmark, Troms (Norway)
Quebec	Quebec (Canada)	Azores	Açores (Portugal)
Western Canada	British Columbia, Alberta (Canada)	Madeira	Madeira (Portugal)
Bornholm	Bornholm (Denmark)	Andalucia	Andalucia (Spain)
Faroe Islands	Faroe Islands (Denmark)	Aragon	Aragon (Spain)
Jutland	Jyllands <i>område</i> minus Sydjylland (Denmark)	Asturias	Asturias (Spain)
North Schleswig	Sydjylland (Denmark)	Balearic Isles	Baleares (Spain)
Aland	Åland/Ahvenanmaa (Finland)	Canary Islands	Canárias (Spain)
Finnish Lappland	Lappi (Finland)	Catalonia	Catalunya (Spain)
Alsace-Lorraine	Alsace, Lorraine (France)	Ceuta and Melilla	Ceuta y Melilla (Spain)
Brittany	Bretagne (France)	Euskadi	País Vasco (Spain)
Burgundy	Bourgogne (France)	Galicia	Galicia (Spain)
Corsica	Corse (France)	Navarre	Navarre (Spain)
West Flanders	Nord <i>département</i> (France)	Valencia	Valencia (Spain)
Normandy	Basse-Normandie, Haute-Normandie (France)	Gotland	Gotland (Sweden)
Occitania	Auvergne, Aquitaine (minus Pyrénées-Atlantiques), Languedoc (minus Pyrénées-Orientales), Midi-Pyrénées, Limousin, Provence-Alpes-Côte d'Azur, Rhône-Alpes (minus Savoie and Haute-Savoie) (France)	Jamtland	Jämtland (Sweden)
Pays Basque Nord	Pyrénées-Atlantiques <i>département</i> (France)	Central Lappland	Norrbotnen (Sweden)
Roussillon	Pyrénées-Orientales <i>département</i> (France)	Scania	Skåne, Halland, Blekinge (Sweden)
Savoy	Savoie, Haute-Savoie <i>départements</i> (France)	Graubünden	Graubünden (Switzerland)
Baden-Württemberg	Baden-Württemberg (Germany)	Ticino	Ticino (Switzerland)
Bavaria	Bayern (Germany)	Northern Ireland	Northern Ireland (UK)
Friesland	Nordfriesland, Emden, Ammerland, Aurich, Wilhelmshaven, Friesland, Oldenburg Landkreis <i>kreisen</i> (Germany)	Scotland	Scotland (UK)
Saarland	Saarland (Germany)	Wales	Wales (UK)
Schleswig-Holstein	Schleswig-Holstein minus Nordfriesland (Germany)	Alaska	Alaska (USA)
Lusatia	Cottbus, Spree-Neisse, Niederschlesischer Oberlausitz, Hoyerswerda, Kamenz <i>kreisen</i> (Germany)	American Midwest	Kansas, Nebraska, Iowa, South Dakota, North Dakota, Minnesota, Missouri (USA)
Saxony	Sachsen minus Niederschlesischer Oberlausitz, Hoyerswerda, Kamenz (Germany)	American South	Virginia, West Virginia, Tennessee, North Carolina, South Carolina, Georgia, Florida, Alabama, Mississippi, Arkansas, Louisiana, Texas, Oklahoma (USA)
Epirus	Ipeiros (Greece)	American West	Washington, Oregon, California, Idaho, Montana, Wyoming, Colorado, Utah, Nevada, Arizona, New Mexico (USA)
Crete	Kriti (Greece)	Hawaii	Hawaii (USA)
Macedonia	Anatoliki Makedonia kai Thraki, Kentriki Makedonia, Dytiki Makedonia (Greece)	New England	Connecticut, Maine, Massachusetts, Vermont, New Hampshire, Rhode Island (USA)
Aegean Isles	Nisia Aigaiou (Greece)	Puerto Rico	Puerto Rico (USA)
Thessaly	Thessalia (Greece)		

Note: Only regions for which sufficient data existed as to include them in the regressions are presented. Capital regions not listed.

Definitions of Moderate and Radical Secessionist Parties Used in Analysis

“MODERATE”	“RADICAL”
Western Canada Concept (Western & Midwestern Canada)	Western Independence Party (Western & Midwestern Canada)
Alliance Démocratique du Québec (Quebec)	Parti Québécois (Quebec)
Folkeflokken (Faroe Islands; through 1994)	Bloc Québécois (Quebec)
Selvstyrepartiet (Faroe Islands; through 1994)	Diverse independentists (Quebec)
Diverse autonomists (Alsace)	Republikanske Parti (Faroe Islands)
Union Democratique Bretonne (Brittany)	Folkeflokken (Faroe Islands; after 1994)
MPA (Corsica)	Selvstyrepartiet (Faroe Islands; after 1994)
Diverse autonomists (Corsica)	Parti pour l'Organization d'une Bretagne Libre (Brittany)
Mouvement Social Occitan (Occitania)	UPC (Corsica)
Diverse autonomists (Occitania)	Corsica Viva (Corsica)
Partit Occitan (Occitania)	Corsica Naziune (Corsica)
Esquerra Catalana dels Treballadores (Roussillon)	A Manca Naziunale (Corsica)
Diverse autonomists (Roussillon)	Rinnovu Naziunale (Corsica)
Unitat Catalana (Roussillon)	Abertzaleen Batasuna (Pays Basque Nord)
Mouvement Savoie (Savoy)	Esquerra Republicana de Catalunya (Roussillon)
Bayernpartei (Bavaria)	Ligue Savoisiennne (Savoy)
Liga Veneta (Padania)	Partidu Independentistu Sardu & variations (Sardinia)
Lega Lombarda (Padania)	Noi Siciliani – FNS (Sicily)
Lega Nord (Padania)	Movimento per l'Indipendenza del Territorio Libero di Trieste (Trieste)
Partito Sardo d'Azione (Sardinia)	Nacion Andaluz (Andalucia)
Lega Sard (Sardinia)	PSM (Balearic Isles)
MAS (Sardinia)	Esquerra Republicana de Catalunya (Balearic Isles)
SiE (Sardinia)	Estat Catala (Balearic Isles)
Sardigna Nazione (Sardinia)	P.N. Canario (Canary Islands)
Mesa Sardos Liberos (Sardinia)	CNC-FREPIC-ANAWAC/CCI (Canary Islands)
Partido Democratico do Atlantico (Azores)	Esquerra Republicana de Catalunya (Catalonia)
Partido Democratico do Atlantico (Madeira)	Herri Batasuna/Euskal Herritarrok (Euskadi)
PPC (Canary Islands)	Euskadiko Ezkerra (Euskadi; through 1984)
UPC (Canary Islands)	Eusko Alkartasuna (Euskadi)
CCNC (Canary Islands)	Frente Popular Galego (Galicia)
AC-INC (Canary Islands)	Herri Batasuna/Euskal Herritarrok (Navarre)
UNI (Canary Islands)	Euskadiko Ezkerra (Navarre; through 1984)
Convergencia i Unio (Catalonia)	Eusko Alkartasuna (Navarre)
Euskadiko Ezkerra (Euskadi; after 1984)	Esquerra Nacional Valenciana (Valencia)
Partido Nacionalista Vasco (Euskadi)	Esquerra Republicana de Catalunya (Valencia)
BNPG/BNG/Bloque PSG (Galicia; through 1989)	Estat Catala (Valencia)
Partido Nacionalista Vasco (Navarre)	Scottish National Party (Scotland)
Euskadiko Ezkerra (Navarre; after 1984)	Alaska Independence Party (Alaska)
BLOC (Valencia)	Partido Independentista Puertorriqueño (Puerto Rico)
Diverse Catalan nationalists (Valencia)	Vlaams Blok (Flanders)
Plaid Cymru/Party of Wales (Wales)	
Libertarian Party of Alaska (Alaska)	
Fronte Autonomista (Padania)	
Volksunie (Flanders; after 1991)	

Data Sources

Provincial boundaries were derived from *The Times Atlas of the World, Tenth Comprehensive Edition* and maintained consistently throughout entire time period.

Variable: *VOTE & VOTE2; LAGVOTE & LAGVOTE2*

Description: Secessionist party vote share

Units: Percentages (0-100 scale)

Sources: Coded by the author from Elections Canada website; Canadian parliament website; Elections Quebec website; Elections Nova Scotia website; Elections PEI website; Elections New Brunswick website; Elections Manitoba website; Elections Saskatchewan website; Elections Alberta website; Elections British Columbia website; Frank B. Feigert, *Canada Votes: 1935-1988* (Durham, N.C.: Duke University Press, 1989); *Elecciones Generales 1989* (Barcelona: Department de Governacià, 1989); *Elecciones Generales, 1982: Congreso* (Madrid: Ministerio del Interior, 1982); *Elecciones Generales, 1979: Congreso* (Madrid: Ministerio del Interior, 1979); <http://www.eleweb.net>; *Elezioni della Camera dei Deputati e del Senato della Repubblica: 14 Giugno 1987* (Rome: Istituto Centrale di Statistica, 1989); *Elezioni della Camera dei Deputati e del Senato della Repubblica, 26 Giugno 1983: Dati Sommari* (Rome, Istituto Centrale di Statistica, 1984); *Elezioni Politiche 1979* (Rome: Istituto Poligrafico e Zecca dello Stato, 1982); *Elezioni Politiche, 21 Aprile 1996* (Rome: Istituto Poligrafico e Zecca dello Stato, 1996); *Risultati delle Elezioni della Camera dei Deputati del 27 e 28 Marzo 1994* (Rome: Camera dei Deputati, 1994); *Risultati delle Elezioni della Camera dei Deputati del 5 Aprile 1992* (Rome: Camera dei Deputati, 1992); Ministero dell'Interno (Italy) website; statistical yearbook of Italy; *Élections Législatives, Résultats des Élections du 24 Novembre 1991: Renouvellement Intégral de la Chambre des Représentants* (Belgium: Directie Verkiezingen en Bevolking, 1991); *Élections Législatives, Résultats des Élections du 13 Décembre 1987: Renouvellement Intégral de la Chambre des Représentants* (Belgium: Directie Verkiezingen en Bevolking, 1988); *Élections Législatives, Résultats des Élections du 13 Octobre 1985: Renouvellement Intégral de la Chambre des Représentants* (Belgium: Directie Verkiezingen en Bevolking, 1986); *Élections Législatives, Résultats des Élections du 8 Novembre 1981: Renouvellement Intégral de la Chambre et du Sénat* (Belgium: ?, 1981); *Élections Législatives, Résultats des Élections du 17 Décembre 1978: Renouvellement Intégral de la Chambre et du Sénat* (Belgium: ?, 1978); statistical yearbook of Belgium; <http://www.vub.ac.be/POLI/elections/Browser.html>; *Statistiske Efterretninger: Befolkning og Valg* (Copenhagen: Danmarks statistik, 1968-1988); *Folketingsvalget den 11 Marts 1998: Danmark,*

Færoerne, Grønland (Copenhagen, Indenrigsministeriet, 1999); *Folketingsvalget den 21 September 1994: Danmark, Færoerne, Grønland* (Copenhagen, Indenrigsministeriet, 1995); *Folketingsvalget den 12 December 1990: Danmark, Færoerne, Grønland* (Copenhagen, Indenrigsministeriet, 1991); statistical yearbook of Denmark; Colin Rallings, *British Electoral Facts, 1832-1999* (Aldershot, Ashgate, 2000); Robert Garner and Richard Kelly, *British Political Parties Today* (Manchester, UK: Manchester University Press, 1998); <http://www.election.demon.co.uk>; <http://www.electionworld.org>; *Elections in the Northern Territory* (Klemzig: Polity Publications, 1994); Colin A. Hughes, *A Handbook of Australian Government and Politics, 1975-1984* (Canberra: Australian National University Press, 1986); Australian Electoral Commission, *Election Statistics. Result of count, Senate and House of Representatives* (Canberra: Australian Government Publications Service, 1990, 1993); *Election Statistics: Australian Capital Territory, Northern Territory* (Canberra: Australian Government Publications Service, 1975-1990); *Election Statistics: Tasmania* (Canberra: Australian Government Publications Service, 1975-1990); *Election Statistics: Western Australia* (Canberra: Australian Government Publications Service, 1975-1990); *Election Statistics: Queensland* (Canberra: Australian Government Publications Service, 1975-1990); James Jupp, *Party Politics, Australia, 1966-1981* (Sydney: Allen and Unwin, 1982); statistical yearbook of Australia; Adam Carr's Electoral Archive website, <http://psephos.adam-carr.net/index41.html>; <http://www.earthsci.unimelb.edu.au/~blair//elections/eleclist.html>; *Stortingsvalget* (Christiana: I Kommission hos H. Aschehoug, 1977-1993); statistical yearbook of Norway; <http://www.ssb.no/stortingsvalg/histtab>; Swiss Federal Statistical Office website; *Kansanedustajain Vaalit/Parliamentary Elections* (Helsinki: Tilastokeskus, 1966-1991); statistical yearbook of Finland; Statistisches Bundesamt, *Wahl zum Deutschen Bundestag* (Stuttgart: W. Kohlhammer, 1969-1990); statistical yearbook of Germany; Bundestag Elections website, <http://www.destatis.de/wahlen>; *Apotelesmata ton vouleitikon eklogon* (Athens: Ethnikon Typographers, 1977-1985); *Hoi ekloges tes 10es Oktovriou 1993* (Athens: Ekdoseis "To Pontiki," 1994); *Hoi ekloges 5es Novembriou '89* (Athens: Ekdoseis "To Pontiki," 1990); Steve R. Reed, *Japan Election Data: The House of Representatives, 1947-1990* (Ann Arbor: Center for Japanese Studies, University of Michigan, 1992); *Statistiek der Verkiezingen. Tweede Kamer der Staten-Generaal* (The Hague: Central Bureau voor de Statistiek, 1977-1994); *Statistiek der Verkiezingen. Provinciale Staten* (The Hague: Staatsdrukkerij- en Uitgeverijbedrijf, 1978-1999); *Eleição para a Assembleia da República, 1976 : resultados por freguesias, concelhos e distritos, comparados com os de 1975* (Lisbon: Imprensa Nacional - Casa da

Moeda, 1976); *1980 eleição da Assembleia da República : escrutínio provisório, resultados por freguesia, resultados por concelho e distrito/região autónoma, comparados com os de 1979* (Lisbon: STAPE, 1980); *1983 eleição da Assembleia da República : escrutínio provisório, resultados por freguesia, resultados por concelho e distrito/região autónoma, comparados com os de 1980* (Lisbon: STAPE, 1983); *1985 eleição da Assembleia da República : escrutínio provisório, resultados por freguesia, resultados por concelho e distrito/região autónoma, comparados com os de 1983* (Lisbon: STAPE, 1985); *Eleição da Assembleia da República, 1987 : escrutínio provisório, resultados por freguesia, resultados por concelhos, distritos e regiões autónomas, comparados com os de 1985* (Lisbon: STAPE, 1987); *Eleição da Assembleia da República, 1991 : escrutínio provisório, resultados por freguesia, resultados por concelhos, distritos e regiões autónomas, comparados com os de 1987* (Lisbon: STAPE, 1991); STAPE website, <http://www.stape.pt>; *Allmänna valen* (Stockholm: Statistiska Centralbyrån, 1976-1991); statistical yearbook of Sweden; *Nationalratswahlen* (Bern: Eidgenössisches Amt, 1975-1995); statistical yearbook of Switzerland; <http://eleccionespuertorico.org>; <http://clerkweb.house.gov/elections/elections.htm>; Inter-university Consortium for Political and Social Research, *State Legislative Election Returns in the United States, 1968-1989* [Computer file], Fifth ICPSR ed. (Ann Arbor, MI: Inter-university Consortium for Political and Social Research [producer and distributor], 1992); Christian Collet, *State Legislative Election Candidate and Constituency Data, 1993-1994* [Computer file], ICPSR version (Irvine, CA: University of California, Irvine [producer], 1996; Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 1997); Alaska Secretary of State website; Hawaii Secretary of State website.

Variable: *LNVOTE & LNVOTE2; LNLAG & LNLAG2*

Description: Logarithmic transformation of above

Units: Logarithmic transformation of percentages

Sources: See above

Variable: *ENEP*

Description: Effective number of electoral parties in province/region at last election

Units: See text

Sources: See above

Variable: *IDEOL & LEFT*

Description: Absolute value of vote for right at last election in province minus vote for right at last election in whole country

(*IDEOL*); vote for right at last election in province minus vote for right at last election in whole country (*LEFT*)

Units: Percentages (0-100 scale)

Sources: See above

Variable: *PROVELEC/REGELEC*

Description: Provincial/regional election=1; national election=0

Units: Dummy variable

Sources: See above

Variable: *REP*

Description: Percentage provincial/regional representation in legislature (lower house)

Units: Percentages (0-100 scale)

Sources: See above

Variable: *POP*

Description: Provincial/regional population

Units: Thousands of persons

Sources: Eurostat REGIO dataset; Economagic.com website; statistical yearbooks of Australia, Belgium, Denmark, Finland, Germany, Greece, Italy, Japan, Netherlands, Norway, Portugal, Spain, Catalonia, Sweden, Switzerland, United Kingdom; U.S. Census Bureau website; Statistics Canada website, <http://www.statcan.ca>; <http://www.library.uu.nl/wesp/populstat/populhome.html>; Statistics Norway website, <http://www.ssb.no>, and correspondence; http://www.social-science-geis.de/en/social_monitoring/social_indicators/EU_reporting/pdf_files/populregio.pdf; Swiss Federal Statistics Office correspondence.

Variable: *GDP RATIO*

Description: Provincial/Regional GDP per capita divided by national GDP per capita

Units: Ratios

Sources: See above

Variable: *UNEMDIFF*

Description: Provincial/Regional unemployment rate minus national unemployment rate

Units: Percentages (0-100 scale)

Sources: See above; also,

http://alt.seco-admin.ch/db/AMS/Archiv/f_talqkant.asp and Bureau of Labor Statistics website

Variable: *MISIND*

Description: Provincial/Regional unemployment rate plus national inflation rate

Units: Percentages (0-100 scale)

Sources: See above; also, World Development Indicators 2002 CD-ROM

Variable: LANG

Description: Percentage of provincial/regional population speaking a region-specific minority language at latest available date

Units: Percentages (0-100 scale)

¹ See for example *Ethnic Conflict in the Western World*, ed. Milton J. Esman (Ithaca: Cornell University Press, 1975); *National Separatism*, ed. Colin H. Williams (Cardiff: University of Wales Press, 1982); *New Nationalisms of the Developed West: Toward Explanation*, ed. Edward A. Tiryakian and Ronald Rogowski (Boston: Allen and Unwin, 1985).

² See for example Joseph R. Rudolph, Jr. and Robert J. Thompson, "Ethnoterritorial Movements and the Policy Process," *Comparative Politics* (April 1985): 291-311. The revisionist "wave" was a small one: most of the scholars who had studied secessionism in the 1970s simply dropped it in the 1980s, rather than re-examining the hypotheses put forward during the previous decade.

³ See for example Kisangani N. Emizet and Vicki L. Hesli, "The Disposition to Secede: An Analysis of the Soviet Case," *Comparative Political Studies* 27, 4 (1995): 493-536; Daniel Treisman, *After the Deluge: Regional Crises and Political Consolidation in Russia* (Ann Arbor: University of Michigan Press, 1999).

⁴ Examples of small secessionist movements that have emerged in the last 20 years include: the Alaska Independence Party, the Ligue Savoisiennne/Savoyan League, Nación Andaluz/Andalusian Nation, the Southern Party, Mec Vannin/Sons of Man, Elsass-Lothringen Nationalforum/Alsace-Lorraine National Forum, and Parti pour l'Organisation d'un Bretagne Libre/Party for the Organization of a Free Brittany, just to name a few. Independence movements that have not yet organized for electoral contests exist in Scania and Jaemtland (Sweden), Aaland (Finland), Hawaii (U.S.), Jura (Switzerland), Valle d'Aoste and Trieste (Italy), Okinawa (Japan), etc. The "Homelands" website purports to list secessionist and autonomist movements all over the world: many of these movements seem to be nothing more than websites, however. See <http://www.visi.com/~homelands>.

⁵ R. William Ayres and Stephen M. Saideman, "Is Separatism as Contagious as the Common Cold or as Cancer? Testing the International and Domestic Determinants of Secessionism," *Nationalism and Ethnic Politics* 6, 3 (2000); R. William Ayres and Stephen M. Saideman, "Determining the Causes of Irridentism: Logit Analyses of Minorities at Risk Data from the 1980s and 1990s," *The Journal of Politics* 62, 4 (2000): 1126-44.

⁶ "Padania" refers to the regions of Italy north of the Po River, not including the small German-, French-, and Slovenian-speaking areas. The term is used especially by the LN and other secessionists from the north.

⁷ The countries included in the analysis were: Australia, Belgium, Canada, Denmark, Finland, France, Germany, Greece, Italy, Japan, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, the U.K., and the U.S.A. The data panels are the first-tier territorial units within these countries (e.g., states in Australia, regions in Belgium, provinces in Canada). These territorial units are all called "provinces" in the text. The term "region" used in the discussion of the hypotheses refers to the territorial entity for which secession is sought as a single unit. "Regions" do not always coincide with "provinces." The Methodological Strategies section deals with this discrepancy.

⁸ For example, Catalan is coded as the regional language for Catalonia, Valencia, and the Balearic Isles, which are separate territorial units but all part of a greater Catalan-speaking region.

⁹ Other instructive examples of such enclaves are Valle d'Aoste in Italy and Aaland in Finland.

¹⁰ A "between" case might be that of Québécois, who speak a language spoken in many other countries, but none of them adjacent to Quebec. French would do not die if the Québécois were assimilated, but it would almost certainly die in North America. The "between" cases are counted with the Welsh and Basques in the regression analyses.

Sources: Ethnologue.com;
http://www.helsinki.fi/~tasalmin/europe_report.html

Variable: *LANGADJ*

Description: =1 if region/province with own language is adjacent to independent country in which that language is an official one, =0 otherwise

Units: Dummy variable

Sources: Coded by author

Variable: *INDEP*

Description: =1 if region/province has been independent or served as the capital region of an independent country at any time since 1648, =0 otherwise

Units: Dummy variable

Sources: Coded by the author, occasionally with the assistance of Britannica.com and the Perry-Castañeda historical maps collection, <http://www.lib.utexas.edu/maps/historical/index.html>

¹¹ Both the secession of the American colonies from Great Britain and the secession of the Confederacy from the United States were stimulated largely by political-economic conflicts.

¹² See Donald L. Horowitz, *Ethnic Groups in Conflict* (Berkeley: University of California Press, 1985), p. 258; Donald L. Horowitz, "Patterns of Ethnic Separatism," *Comparative Studies in Society and History* 23 (April 1981): 165-95.

¹³ Jason Sorens, "Globalization, Autonomy, and Secessionism," unpublished manuscript, Department of Political Science, Yale University (May 2002).

¹⁴ Sorens 2002.

¹⁵ World production differs from GDP in that it excludes construction and non-tradeable services.

¹⁶ In the language of economics, provision of most public goods enjoys declining average costs over population.

¹⁷ Horowitz (1985), p. 621.

¹⁸ For U.S. states representation in both the House of Representatives and Senate is used.

¹⁹ Sorens 2002.

²⁰ See Gary Cox, *Making Votes Count*. Crucial here is the fact that electoral rules do not just have a mechanical effect on the translation of votes into seats but also have a psychological or strategic effect, reducing the actual vote shares for parties that are not perceived as having a chance of winning seats.

²¹ It is measured from elections to the lower house of a parliament. Puerto Rico is an exception due to its unusual electoral system that makes it impossible to calculate a single party vote share. For Puerto Rican elections, the vote for Resident Commissioner was used for variables at a countrywide-election observation, while the vote for Governor was used for variables at a provincial-election observation. In countries such as Germany and Italy in which some delegates are elected in plurality constituencies, others through proportional representation (PR), only the PR ballots are used to calculate party vote shares. All the above criteria are also used for the calculation of *IDEOL*, *LEFT*, the dependent variables, and the lagged secessionist vote variables described in the text.

²² Markku Laakso and Rein Taagepera, "'Effective' Number of Parties: A Measure with Application to West Europe," *Comparative Political Studies*, 12 (1979), 3-27.

²³ The variables are differentiated according to election type: the lagged vote variable for a provincial election measures vote at the last provincial election, while the lagged vote variable for a countrywide election measures vote at the last countrywide election. If, however, the provincial election is the first ever (as in Scotland and Wales in 1999), the vote from the last countrywide election is used.

²⁴ See for example Robert W. Jackman and Karin Volpert, "Conditions Favouring Parties of the Extreme Right in Western Europe," *British Journal of Political Science* 26 (1996), 501-21.

Variable: *DIST*

Description: =1 if region/province lacks road connections with rest of country, =0 otherwise

Units: Dummy variable

Sources: Coded by the author

Variable: *GLOB*

Description: World exports divided by world production

Units: Index scale (1990=1.0)

Sources: Calculated by the author from International Financial Statistics 2000 CD-ROM

Variable: *PROVAUTO/REGAUTO & DAUTO/DREGAUTO*

²⁵ In these regressions observations from the United States constituted nearly half of all observations, due to the high number of administrative divisions, frequency of elections, and availability of data. When U.S. observations are removed, results do not differ substantially. U.S. observations did not need to be removed from the “linguistically distinctive” regressions because only Alaska, Hawaii, and Puerto Rico were included in this dataset.

²⁶ If the unemployment variables are dropped and a dummy variable for Spanish observations is included, all the other significant independent variables retain their significance, while the Spanish dummy is highly significant and positive. There is much more secessionism in Spain than the model would otherwise predict.

²⁷ Sorens 2002.

²⁸ The point predictions that follow are not to be taken too seriously without the standard errors; the analysis that follows is intended merely to examine the relative importance of the significant independent variables. Linear interpretations of the variables with standard errors of the sort “X increases Y plus or minus Z units” are not terribly useful here in any case, since the dependent variable is a non-linear transformation of the variable of interest, vote percentages.

²⁹ It is obvious simply from the coefficient estimates how important the dummy variables are in relation to each other. From Table 2, *LANGADJ* is more important than *INDEP*, which is slightly more important than *DIST*, which has a coefficient twice as large as *PROVELEC*. In Table 3 the results are basically reversed: *DIST* most important, *LANGADJ* least, with *INDEP* between.

³⁰ The modestly nonlinear relationship between language and secessionism revealed below makes this particular counterfactual somewhat dubious, however.

³¹ *ENEP*'s coefficients are actually slightly smaller in the limited sample, but this difference may be explained simply by the contraction of the dataset and the loss in degrees of freedom. Variables like *GDPRATIO* see their coefficients contract much more.

³² This result seems to contradict directly “modernization” theories of ethnic conflict, according to which ethnic conflict heightens when groups are mixed together and interacting with each other often.

³³ Poll data cited in “Flemish Questions” website:

http://home.online.no/~vlaenen/flemish_questions/quste41.html.

³⁴ Note: in all historical-regions regressions *REGELEC* is the equivalent of *PROVELEC*, *REGAUTO* is a regional autonomy variable measured slightly differently from *PROVAUTO* and ranging from 0-3, and *DREGAUTO* measures change in *REGAUTO*. *REGAUTO* is measured in the following way: +1 if the region or its mutually exhaustive, geographically defined subunits have an elected executive(s), +0.5 if the region exists as a single political entity for the purpose of regional government elections, +0.5 if the region or its mutually exhaustive subunits are legally superior to geographically lower-level units and are not themselves subject to any geographically higher unit except the central government, +0.5 if regional or mutually exhaustive subregional governments enjoy both administrative and legislative powers, +0.5 if regional or mutually exhaustive subregional governments derive more than 25% of their revenue from own-source taxes.

Description: Level & change of provincial autonomy, 0-4 scale
Units: Step scale
Sources: Coded by the author from OECD, "Managing across Levels of Government" (1997), <http://www1.oecd.org/puma/malg/malg97/toc.htm>; for the Faroe Isles from Benoît Raoulx, *Les Îles Féroé* (Caen Cedex: Centre de Publications de l'Université de Caen, 1992)