I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 15570

Is Secessionism Mostly about Income or Identity? A Global Analysis of 3,003 Subnational Regions

Klaus Desmet
Ignacio Ortuño Ortín
Ömer Özak

SEPTEMBER 2022

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 15570

Is Secessionism Mostly about Income or Identity? A Global Analysis of 3,003 Subnational Regions

Klaus Desmet
SMU

Ignacio Ortuño Ortín
UC3M

Ömer Özak
SMU and IZA

SEPTEMBER 2022

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Is Secessionism Mostly about Income or Identity? A Global Analysis of 3,003 Subnational Regions*

This paper analyzes whether the propensity to secede by subnational regions responds mostly to differences in income per capita or to distinct identities. We explore this question in a quantitative political economy model where people's willingness to finance a public good depends on their income and identity. Using high-resolution economic and linguistic data for the entire globe, we predict the propensity to secede of 3,003 subnational regions in 173 countries. We validate the model-based predictions with data on secessionist movements, state fragility, regional autonomy, and conflict, as well as with an application to the dissolution of the Soviet Union. Counterfactual analysis strongly suggests that identity trumps income in determining a region's propensity to secede. Removing identity differences reduces the average support for secession from 7.5% to 0.6% of the population.

JEL Classification: H77, P00, D70, D74, F02, F52, Z10

Keywords: secessionism, separatism, federalism, conflict, identity, political economy, international relations

Corresponding author:
Ömer Özak
Dept. of Economics
Southern Methodist
University 3300 Dyer St.
Box 0496
Dallas, TX 75275
USA

E-mail: ozak@smu.edu

* We thank Angel Cuevas, Rubén Cuevas and José González Cabañas for help with the Wikipedia data, and Ignacio Ortuño Ortín acknowledges financial support from the Spanish Ministry of Science, Innovation and Universities (Project PID2019-109157RB-IOO).

1 Introduction

Secessionist tensions pose a fundamental challenge for territorial governance across the globe, from Flanders and Catalonia in Europe, to Nagorno-Karabakh and Tibet in Asia, and Oromia and the Western Cape in Africa. In recent decades, referendums on independence have taken place in regions such as Quebec, East Timor, Kurdistan, Somaliland, South Ossetia, South Sudan, Scotland, and Puerto Rico, and pro-independence or regionalist political parties hold significant vote shares in regions such as the Basque Country and Lombardia. Although successful secessions have been uncommon since the breakup of the Soviet Union and Yugoslavia in the early 1990s,[1] there are currently active secessionist movements in 54% of countries and 17% of subnational regions of the world. Some of these regions are relatively wealthy, spurring discontent among their population about subsidizing the rest of the country. Others exhibit a distinct identity, eroding their sense of loyalty to the nation. While there is general agreement that income per capita and identity are fundamental drivers of the demand for secession, their relative importance remains debated.

This paper asks whether secessionism tends to respond mostly to income per capita or identity differences. It tackles this question at an unprecedented global scale, for all first-level administrative regions of the world. Our analysis relies on a quantitative political economy model where people are less willing to contribute to a public good if they are richer or if they have an identity that is different from the rest of the country. Taking the model to the data for the entire world, we predict for 3,003 subnational regions the share of the population that favors secession. Counterfactual analysis that selectively switches off the model's different forces suggests that identity trumps income in determining a region's propensity to secede.

In our model, scale economies, income heterogeneity, and identity differences determine a country's stability. A country's geography is partitioned into subnational regions, each with a population size and an income level. A country's population is partitioned into identity groups that may or may not coincide with the subnational regions. Individuals have preferences over private consumption and public consumption. The more an individual's identity differs from that of the rest of the population, the lower the utility she derives from public consumption. Public goods are financed through a proportional income tax, decided by majority vote. As a result, richer individuals, as well as those with a distinct identity, end up paying a higher tax rate than they prefer. Hence, subnational regions with higher income per capita and a different identity from the rest of the country experience on average greater support for independence. Pushing in the other direction are the scale economies associated with the provision of public goods, as well as within-region identity heterogeneity.

We quantify the model using high-resolution data on population, income per capita and identity for the entire globe. For population, we use data from Landscan for the year 2000; for income per capita, the data come from G-Econ 4.0 and are also for the year 2000; for identity, we rely on a high-resolution database on language use developed by Desmet, Gomes and Ortuño Ortín (2020). Although the theory applies to any dimension of identity, we focus on language for two reasons: language has long been identified as a major identity marker that differentiates populations (Fearon, 2003), and linguistic data provide global coverage at the subnational level. For all countries of the world, we aggregate these high-resolution population, income, and linguistic data up to first-level administrative regions (e.g., U.S. states, German Länder, Russian oblasts, Colombian departments, Kenyan provinces, Japanese prefectures). The model's parameters are then calibrated to match the territorial divisions of the current world map. This calibration is motivated by the observation that, in spite of widespread territorial

[1] The only three successful secessions have been East Timor, Eritrea, and South Sudan.

challenges and conflicts, successful secessions or unions have been rather rare in recent decades.[2]

We use our calibrated model to estimate the share of each region's and each country's population in favor of secession. Within the top-10% group of regions with strongest support for independence, we find regions such as Tibet (China), Aceh (Indonesia), Lombardia (Italy), Okinawa (Japan), Tatarstan (Russia), and Catalonia (Spain). The most unstable countries for each of the six major continents are India, Italy, Papua New Guinea, South Sudan, Guatemala, and Bolivia. The share of the population that wants to secede is largest in South Asia (14%), followed by sub-Saharan Africa (13%), and it is smallest in North America, where less than 1% of the population favors secession.

To validate our calibrated model, we pursue three different strategies. First, we construct a global geographic database of active secessionist movements and analyze how well our model-generated measures of secessionism fit those data. More specifically, starting with a comprehensive Wikipedia list of 2,529 active secessionist movements, we identify the first-level administrative regions associated with each movement. To get a measure of the importance of each secessionist movement, we use the number of Wikipedia page views in all languages over a five-year period. We find that the predictions of the model align well with actual secessionist activity across the world, both at the country and the regional level. Second, we compare our model predictions to data on the vulnerability of states to collapse, the autonomy of regional governments, and the intensity of conflict within countries. Reassuringly, in all cases there is a strong association with our measures of demand for secession. Third, using data of the end of the 1980s, we show that our calibrated model is able to account for the breakup of the Soviet Union. The model predicts not only that the country was unstable, but also that the Soviet republics most in favor of secession were the first ones to declare independence, and those least in favor of secession were the last ones to leave the union. In that sense, while we calibrate the model to the current world map, we are able to explain the perhaps most dramatic territorial change of the past half-century.

We then use counterfactual analysis to investigate the relative importance of income per capita and identity in determining the propensity to secede. First, we show that removing income per capita differences across subnational regions within countries has a limited effect on the demand for secession, whereas removing identity differences makes support for secession vanish. More specifically, starting from a baseline average regional support for secession of 7.5%, we see a drop by 6.9 percentage points in the absence of identity differences and no decline in the absence of income per capita differences. This suggests that identity is the main driver of the demand for separatism, whereas income per capita plays only a minor role. Second, we further assess how sensitive secessionism is to changes in income per capita. Focusing on regions where at least 10% of the population supports independence, we find that income per capita would on average have to drop by 43% for its residents to give up on secessionism. That is, secessionism only weakens when separatist regions fall far behind economically. This points, again, to economic forces being of limited importance in shaping the drive for independence. Third, we analyze the effect of removing within-region identity differences, without changing the overall identity heterogeneity of the country. This leads to an almost tripling of the support for secessionism at the regional level: from 7.5% in the baseline to 20.0%. Hence, if subnational regions had a homogeneous identity, the model predicts that countries would become much more unstable. Once again, this shows that identity is an important force in shaping secessionism.

To further explore the robustness of our findings, we also consider an alternative calibration that targets public

[2]Apart from the three successful secessions mentioned before, in the last thirty years the world has witnessed only two effective unions (West and East Germany, and North and South Yemen). The relative stability of the world map may reflect that secessionist tensions, though common, often fail to garner large majorities. Sorens (2008) finds a strong presence of regionalist parties in many advanced democracies, but they rarely surpass the 50% vote share. Similarly, the well-known referenda on independence in Quebec (1995) and Scotland (2014) did not succeed.

spending on defense and the regionalist vote share in a restricted set of developed countries. The counterfactual analysis under this alternative calibration confirms our main finding: identity is the main driver of secessionism. Removing income per capita differences has essentially no effect on the demand for secession, whereas removing identity differences lowers the support for secession from an average 6.4% of a region's population to 0.5% in this alternative calibration.

Taken together, these results strongly suggest that identity, more than income, is key to understanding secessionism. Hence, while economic policies might be of some help in staving off secessionist threats, creating a common identity that cuts across existing ethnolinguistic cleavages is more likely to lead to territorial stability. Indeed, such policies were central to European nation-building efforts during the nineteenth century (Alesina, Giuliano and Reich, 2021). In Italy, for example, only a small proportion of its population spoke Italian at the time of unification in the mid-19th century. The adoption of a uniform and common language, based on the Florentine dialect, was viewed as an essential element in creating a unified Italian state. Similarly, at the time of the French Revolution, only 10% of the population spoke French fluently. The introduction of mass education under the *Loi Guizot* of 1833 has been credited with being the "ultimate acculturation process that made the French people French" (Weber, 1979).

Our paper is related to the theoretical, empirical, and quantitative literature on the stability of countries. Theoretical work on the size and number of nations has focused on the tradeoff between the benefits of scale economies and the costs of both preference heterogeneity (Alesina and Spolaore, 1997, 2003) and disagreements over fiscal redistribution (Bolton and Roland, 1997). This tradeoff is present in our framework as well: larger countries benefit from scale economies in the provision of public goods, but tend to suffer from greater regional differences in income per capita and identity.

The empirical literature on the relative importance of identity and income as drivers of secessionism is inconclusive. Focusing on advanced economies, Sorens (2008) finds that secessionist parties are stronger in richer regions, whereas non-secessionist regionalist parties are stronger in areas that strive for cultural autonomy.[3] In an analysis of Western Europe, Álvarez Pereira, Portos and Vourdas (2018) emphasize the interaction between economic and cultural variables: support for autonomy and secession is greater in richer regions, but only to the extent that they are culturally differentiated. Suesse (2019) studies separatist protests in 183 provinces of the Soviet Union between 1987 and 1992, and finds evidence of more protests in larger territories that are ethnically distant from the rest of the country. Compared to our work, a shortcoming of these studies is their limited geographical scope and their omission of intra-regional identity heterogeneity.

There are a few quantitative papers on country stability, but they focus exclusively on Europe. Desmet et al. (2011) use a similar model to ours and quantify it to Western Europe, using genetic distances as a proxy for regional distinctiveness. Vanschoonbeek (2020) also analyzes how secession-prone different European regions are, but uses political distinctiveness in voting patterns as a proxy for regional differences. In contrast to these papers, we consider the entire globe, and hence include the many areas outside Europe where secessionist movements are strong. This global approach also allows us to make more general statements about the relative importance of income and identity as drivers of secessionism. Moreover, our analysis highlights the importance of within-region identity heterogeneity, a force that had remained unexplored by these previous papers. We find that the lack of a homogeneous identity in many subnational regions is a key impediment to secessionism.

Rather than focusing on scale economies coming from public goods, others have considered market access

[3]See also Sambanis and Milanovic (2014) who study 48 decentralized economies and conclude that richer regions are more likely to demand greater autonomy.

through trade as the main advantage of being part of a larger union. Dhingra et al. (2017) use a quantitative trade model and estimate static losses from Brexit to UK living standards on the order of 1-3%. Comerford and Rodríguez Mora (2019) also rely on a quantitative trade model to analyze the effect of Scotland, Catalonia and the Basque Country becoming independent nations, and estimate losses in real GDP ranging from 8.5 percent in Scotland to 16 percent in the Basque Country.[4] For the purpose of studying nation stability, we prefer to focus on a model with public goods, because market access through trade is in today's world no longer tightly related to country borders. While in the nineteenth century trade access was mostly achieved by removing country borders through empire building, in the post-WWII period gaining market access has mostly been about signing trade agreements between countries (Gancia, Ponzetto and Ventura, 2021). Country borders are now less of an impediment to trade, as reflected by smaller countries being more open to international markets (Alesina, Spolaore and Wacziarg, 2000).

This paper is also related to the literature on the role of identity in economics and politics. Since the pioneering work of Akerlof and Kranton (2000), there has been a growing realization that social identities affect economic outcomes. An individual's sense of self is related to the social groups she identifies with. Because social groups are associated with norms that determine preferences, it is tempting for governments of nations to influence and shape social identities. The efforts at nation-building through linguistic homogenization are a case in point, but this is not the only way for nations to form a common identity. For example, in the context of pre-revolutionary France, regions where the local feudal regime had been replaced by national fiscal institutions were more likely to identify with the French state (Johnson, 2019). In the same way that national policies can promote the cohesion of the nation-state, regional policies can enhance separatist sentiment. For example, Clots-Figueras and Masella (2013) show that individuals who were more exposed to teaching in Catalan are more likely to feel more Catalan than Spanish, independently of whether their parents are of Catalan origin.[5] Our work shares with this literature the insight that an individual's social identity matters for her utility, hence affecting her choices.

The rest of the paper is organized as follows. Section 2 proposes a simple model of a region's propensity to secede; Section 3 describes the data and the calibration; Section 4 reports the model-predicted propensity to secede for all first-level administrative regions of the world; Section 5 assesses the model's performance; Section 6 conducts counterfactual analysis to determine the relative importance of identity and income in driving the demand for secession; and Section 7 concludes.

2 A Simple Theory of the Propensity to Secede

In this section we introduce a political economy model that explores the relative importance of economic and identity differences as drivers of the demand for secession. Our simple framework is based on Desmet et al. (2011), and captures the trade-off between scale economies and heterogeneity in determining the stability of a country. The scale economies come from the existence of public goods that are financed through taxes, and the heterogeneity comes from differences in income per capita and identity.

[4] In related work, Gurevich et al. (2020) explore the role of language as an internal trade barrier. In particular, they estimate a gain in Canadian GDP by 1.22 percent if Quebec were to become fully bilingual in English.

[5] An individual's social identity is of course not limited to the region or the nation she belongs to. Race, gender and religion also constitute important identity cleavages (Desmet and Wacziarg, 2021). In recent decades, these cleavages have contributed to the rise of identity politics in many advanced democracies (Fukuyama, 2018; Bonomi, Gennaioli and Tabellini, 2021; Gethin, Martínez-Toledano and Piketty, 2021).

2.1 Setup

Countries, subnational regions and identity groups. A country C is geographically partitioned into $R(C)$ subnational regions, indexed by r. There are $N(C)$ individuals, partitioned into L identity groups, indexed by ℓ or k. Denote by $N(\ell, r, C)$ the number of individuals of group ℓ living in region r of country C. Income per capita differs across regions but not within regions.[6] Denote by $y(r, C)$ the income per capita of an individual who resides in region r of country C, and by $Y(C)$ the total income in country C.

We think of regions as administrative units, and of identity groups as collections of individuals defined by a common identity, such as language, ethnicity or religion. Often there is some overlap between regions and identity groups, but typically that overlap is not perfect. To give an example, take Canada and consider language as a group's identity marker. Most people in Quebec are French-speaking, and most people in New Brunswick are English-speaking, but both provinces have sizable minorities speaking the other language.

Preferences over private and public goods. An individual of group ℓ residing in region r and country C has quasi-linear preferences over private consumption $x(\ell, r, C)$ and public consumption $G(C)$ of the form

$$u(x(r, C), G(C), S(\ell, C)) = x(\ell, r, C) + \alpha S(\ell, C)^{\delta} G(C)^{\beta} \tag{1}$$

where $\alpha > 0$ and $\beta, \delta \in [0, 1]$, and $S(\ell, C)$ is the share of people in country C that belong to identity group ℓ. The idea is that an individual derives more utility from the public good, the stronger she feels identified with the country. To use the same example as before, to an English-speaking individual in Canada, public consumption is more valuable the greater the share of Canadians that are English-speaking. This is consistent with Li (2010) who in a cross-country study finds that individuals from the ethnic majority group have higher tax morale, and so do individuals who more strongly identify with their nations. What do the different parameters in the utility function capture? A higher value of α implies giving more importance to the public good; a higher value of β reduces the concavity of the utility derived from the public good; and a higher value of δ makes utility more sensitive to small changes in the degree of identification with the nation.

2.2 Optimal Tax Rate

Public consumption $G(C)$ is financed by a proportional tax rate $\tau(C)$ on income, so $G(C) = \tau(C)Y(C)$. As a result, private consumption $x(\ell, r, C) = (1 - \tau(C))y(r, C)$. Hence, the indirect utility of an individual of identity group ℓ who resides in region r and country C is

$$v(y(r, C), S(\ell, C), \tau(C), Y(C)) = y(r, C)(1 - \tau(C)) + \alpha S(\ell, C)^{\delta} (\tau(C)Y(C))^{\beta}. \tag{2}$$

The tax rate $\tau(C)$ is decided by majority vote.[7]

[6] Allowing for income differences within regions is a straightforward extension. However, because of data limitations on within-region income inequality, we do not explore this possibility here.

[7] In non-democratic countries, where people do not vote, regimes are arguably more likely to last if they maximize the total well-being of the population. In our framework, this coincides with implementing the preferred policy of the mean voter. In the quantification of our model, there is in practice no important difference in terms of country stability between implementing the ideal policy of the mean or the median voter.

Preferred individual tax rate. Maximizing (2) with respect to $\tau(C)$ yields the preferred tax rate of an individual of group ℓ who resides in region r in country C:

$$\tau(\ell, r, C) = \left(\frac{\alpha\beta S(\ell, C)^{\delta} \left(Y(C)\right)^{\beta}}{y(r, C)} \right)^{\frac{1}{1-\beta}}. \tag{3}$$

An individual prefers lower taxes if either her income per capita is higher or her identification with the country is weaker. We therefore say that an individual's preferred tax rate is decreasing in her income-to-identity ratio, denoted by:

$$\tilde{y}(\ell, r, C) = \frac{y(r, C)}{S(\ell, C)^{\delta}} \tag{4}$$

so that

$$\tau(\ell, r, C) = \left(\frac{\alpha\beta (Y(C))^{\beta}}{\tilde{y}(\ell, r, C)} \right)^{\frac{1}{1-\beta}}. \tag{5}$$

Country tax rate. The decision by majority vote then implies that a country's tax rate is decided by the individual with the median income-to-identity ratio, $\tilde{y}_m(C, \delta)$,[8] so

$$\tau(C) = \left(\frac{\alpha\beta (Y(C))^{\beta}}{\tilde{y}_m(C, \delta)} \right)^{\frac{1}{1-\beta}}. \tag{6}$$

The quasi-linear preference structure implies that countries with the same population size spend the same on the public good, independently of their income per capita. This captures the idea that public spending, such as defense and government services, tends to scale up with the size of a country. By the same token, this implies that tax rates are larger in poorer nations than in richer nations. If in a parametrized version of the model this would lead to excessively high taxes in the poorest countries, we could simply assume that in low-income countries individuals attach a smaller weight to the public good. As we will theoretically show later, doing so is irrelevant in the context of our model: reducing the importance of the public good by lowering the value of α in the utility function has no effect on the share of the population that supports secession.

2.3 Propensity of Subnational Regions to Secede

To determine the instability of a country, we need to understand whether its residents prefer to remain in the union or secede. If a region r secedes from country C to form its own country, an individual of group ℓ in that region faces the same income per capita as before, but a different degree of identification, $S(\ell, r)$, a different tax rate, $\tau(r)$, and a different total income, $Y(r)$.[9]

Region tax rate. The indirect utility of an individual of group ℓ residing in independent region r is

$$v(y(r, r), S(\ell, r), \tau(r), Y(r)) = y(r, r)(1 - \tau(r)) + \alpha S(\ell, r)^{\delta} (\tau(r) Y(r))^{\beta}. \tag{7}$$

[8] Because the income-to-identity ratio (4) depends on δ, we explicitly include δ when denoting the median income-to-identity ratio.

[9] In case of secession, a region's income per capita might of course vary, but it is unclear in which direction. For example, market fragmentation may lower income per capita, whereas better governance may increase income per capita.

The corresponding preferred tax rate of that individual if her region r were independent is then

$$\tau(\ell, r, r) = \left(\frac{\alpha\beta(Y(r))^{\beta}}{\tilde{y}(\ell, r, r)} \right)^{\frac{1}{1-\beta}} \tag{8}$$

where $\tilde{y}(\ell, r, r) = y(r, r)/S(\ell, r)^{\delta}$. The tax rate of the independent region r is decided by the individual from r with the median income-to-identity ratio, $\tilde{y}_m(r, \delta)$, so

$$\tau(r) = \left(\frac{\alpha\beta(Y(r))^{\beta}}{\tilde{y}_m(r, \delta)} \right)^{\frac{1}{1-\beta}}. \tag{9}$$

Plugging (9) into (7) then gives the utility of a resident of r who belongs to group ℓ under independence.

Preference for secession. An individual of group ℓ residing in region r prefers her region to be an independent country rather than being part of country C if her utility under independence is greater:

$$v(y(r, r), S(\ell, r), \tau(r), Y(r)) > v(y(r, C), S(\ell, C), \tau(C), Y(C)). \tag{10}$$

We can rewrite this condition as

$$y(r) \left(\left(\frac{\beta(Y(C))^{\beta}}{\tilde{y}_m(C, \delta)} \right)^{\frac{1}{1-\beta}} - \left(\beta \frac{(Y(r))^{\beta}}{\tilde{y}_m(r, \delta)} \right)^{\frac{1}{1-\beta}} \right) + S(\ell, r)^{\delta} \left(\frac{\beta(Y(r))}{\tilde{y}_m(r, \delta)} \right)^{\frac{\beta}{1-\beta}} - S(\ell, C)^{\delta} \left(\frac{\beta(Y(C))}{\tilde{y}_m(C, \delta)} \right)^{\frac{\beta}{1-\beta}} > 0. \tag{11}$$

From this condition we can draw three conclusions. First, the benefits from seceding increase in the group's identification with the region, $S(\ell, r)$, and it decreases with the group's identification with the nation, $S(\ell, C)$. Second, the benefits from seceding increase in a region's income per capita, $y(r)$.[10] Third, whether an individual prefers her region to secede is independent of α. This last result is important for the parametrized version of the model in the sense that our findings do not directly depend on whether we take a narrow view of public goods (which would imply a low value of α) or a a broad view of public goods (which would imply a high value of α). It also means that we could assign a lower value of α to the lowest-income countries, without changing whether an individual supports secession.

Aggregating the preferences for secession across population groups, we can determine for each region r the share of the population that wants to secede, and for each country C the share of the population that prefers to leave the union. This gives us the two measures of instability that we will use in our quantitative exploration of secessionism. One is at the level of subnational regions, and the other at the level of countries.

Share of region's population that favors secession. *For region r in country C, the share of its population who wants to secede from country C is given by*

$$I(r, C) = \sum_{\ell} \mathbb{I}(\ell, r, C) S(\ell, r) \tag{12}$$

[10] Note that when a region's income per capita increases, the fraction $(Y(r))^{\beta}/\tilde{y}_m(r, \delta)$ decreases. This comparative statics result assumes that the country's income and its median voter do not change.

where

$$\mathbb{I}(\ell, r, C) = \begin{cases} 1 & : \quad v(y(r,r), S(\ell,r), \tau(r), Y(r)) > v(y(r,C), S(\ell,C), \tau(C), Y(C)) \\ \\ 0 & : \quad \textit{otherwise} \end{cases}$$

Share of country's population that favors secession. *For country C, the share of its population who wants to secede is given by*

$$I(C) = \sum_{\ell} \sum_{r} \mathbb{I}(\ell, r, C) \frac{N(\ell, r, C)}{N(C)} \tag{13}$$

3 Data, Measurement, and Calibration

To explore the quantitative properties of our model, we need data on identity, population and income at the subnational level.

3.1 Data and Measurement

Subnational regions. Our analysis covers 3,003 first-level administrative regions in 176 countries. To give some examples, first-level administrative regions correspond to states in the U.S., provinces in Canada, regions *(oblasts)* in Ukraine, states in Nigeria, provinces *(províncias)* in Mozambique, regions *(régions)* in France, and states *(Länder)* in Germany. The maps delineating the regions come from GADM version 3.6.

Identity. An individual's identity depends on the language she speaks. Although the theory applies more broadly to any dimension of identity, in the empirics we focus on language. We do so primarily because of the availability of high-quality language data at the subnational level. It is also worth remembering that in many parts of the world language is a good indicator of group identity.

How strongly an individual identifies with her country or region depends on the country's or region's population share that speaks her language. An individual of group ℓ residing in region r of country C has a degree of identification $S(\ell, C)$ with her country and a degree of identification $S(\ell, r)$ with her region. We therefore need to know the share of people who speak each language, not just by country but also by region. To that end, we use spatially disaggregated data on language speakers from a new high-resolution database developed by Desmet, Gomes and Ortuño Ortín (2020). They combine spatial data on population from Landscan with a polygon shapefile of nearly 7,000 languages and a matrix of language speakers by country from the World Language Mapping System, the digitized version of the 17th edition of the Ethnologue. Using these data, they apply an iterative proportional fitting algorithm to allocate language speakers to 5 km by 5 km grid cells for the entire globe.[11] By aggregating these grid-cell data up to the regional level, we obtain the linguistic composition for all first-level administrative regions.

When measuring an individual's linguistic identity, it is not always clear which linguistic groups to use as primitives. Should Picard and Franco-Provençal, two variations of French, be considered as two different language groups or should they be aggregated into French? More generally, does linguistic identity depend on coarse divisions, such as between Romance and Germanic languages, or on fine divisions, such as between Neapolitan and

[11] Cross-validating the algorithm, they find a correlation of 0.80 when comparing their local diversity measures with the ones in Gershman and Rivera (2018) which are based on census data of subnational regions in sub-Saharan Africa.

standard Italian? Since there is no ex-ante good answer to this question, for now we parametrize the coarseness of linguistic divisions and defer a discussion of its value to the model's calibration. More specifically, we follow Desmet, Ortuño Ortín and Wacziarg (2012) and use the language tree of the Ethnologue to construct language groups at 15 different levels of aggregation, $\eta = 1, \cdots, 15$. At the most aggregate level, $\eta = 1$, only the largest language families, such as Indo-European and Niger-Congo, are considered to be different groups, whereas at the most disaggregate level, $\eta = 15$, Picard and Franco-Provençal would be considered distinct from French. This procedure provides us with 15 matrices of the number of language speakers by first-level administrative region, one matrix for each level of linguistic aggregation. An individual's degree of identification then depends on the level of linguistic aggregation. For example, a speaker of Picard would feel more strongly identified with France at high aggregation level 1, where Picard belongs to the same linguistic group as French, than at low aggregation level 15, where the two languages belong to different groups.

The degree of linguistic identification of an individual to her country or region can also be reinterpreted as the linguistic proximity to her country or region. If we assume that an individual of group ℓ has linguistic proximity 1 to anyone speaking her own language and linguistic proximity 0 to anyone speaking a different language, then $S(\ell, C)$ is her average linguistic proximity to the rest of the country and $S(\ell, r)$ is her average linguistic proximity to the rest of the region. Given that linguistic proximity is measured on a 0-1 scale, we can correspondingly define the linguistic distance of an individual to her country and region as, respectively, $1 - S(\ell, C)$ and $1 - S(\ell, r)$.

In the model, the incentive of a subnational region to secede increases with the linguistic distance of its residents to the country, and it decreases with the linguistic distance of its residents to itself. To visualize these incentives, Figure 1 depicts the population-weighted linguistic distance between region r and country C, $\sum_\ell S(\ell, r)(1-S(\ell, C))$. Darker-colored subnational regions are linguistically more distant from the country they belong to, and thus have a stronger incentive to secede. Language groups are based on linguistic aggregation level 15, and thus finely defined. Appendix Figure B.1 shows similar maps for languages defined at a coarser level. In certain regions, such as eastern Africa, southern Europe and northern India, we see a significant decline in linguistic distances of subnational regions to their respective countries when defining languages at a coarser level.

Figure 2 represents the population-weighted linguistic distance between region r and itself, $\sum_\ell S(\ell, r)(1 - S(\ell, r))$. Darker-colored subnational regions are linguistically more distant from themselves, and have thus a weaker incentive to secede. Here as well, languages are finely defined, at aggregation level 15. Appendix Figure B.2 shows similar maps for coarser classifications of languages. There, we see important declines in within-region linguistic distances in Europe and parts of Africa.

Population and income per capita. Landscan provides estimates for population at a resolution of 30″ by 30″. For the year 2000, we aggregate those data up to first-level administrative regions. As for measuring GDP, we start with data with a resolution of 1° by 1° from the G-Econ 4.0 project at Yale University. For each 1° by 1° grid-cell, we assign GDP to 14,400 smaller 30″ by 30″ grid-cells according to their population weights from Landscan. We then sum those smaller grid-cells to obtain estimates of GDP in the year 2000 for all first-level administrative regions. Figure 3 shows the GDP per capita of subnational regions relative to that of the countries they belong to. Darker-colored regions have higher GDP per capita compared to their country, and have thus a stronger incentive to secede.

Figure 1: Linguistic Distance between Subnational Regions and Countries

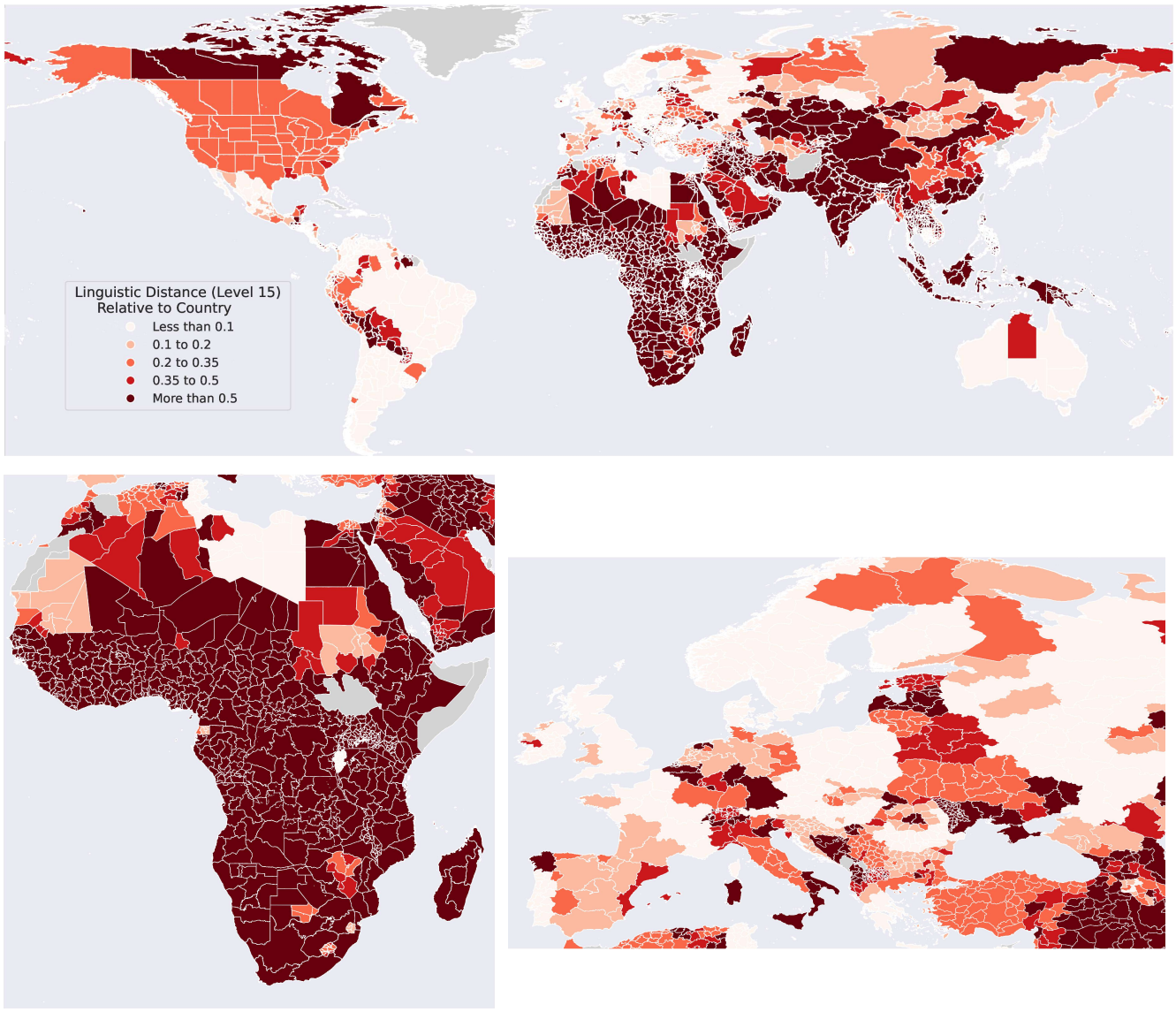


Figure depicts the linguistic distance between each subnational region and the country it belongs to. It is measured as the expected distance between a randomly drawn individual of the subnational region and a randomly drawn individual of the country. Darker colors indicate greater linguistic distances between region and country, and hence a stronger propensity to secede. Languages are defined at its finest level (aggregation level 15).

3.2 Calibration

To numerically assess the model's predictions, we need values for α and β, the parameters that determine the importance and the curvature of the utility derived from public goods, and for δ and η, the parameters that determine the sensitivity of utility to the degree of identification.

Baseline calibration. Our baseline calibration is motivated by the observation that, in spite of widespread territorial tension, the world map is relatively stable. Although there are secessionist problems around the globe,

Figure 2: Within-Subnational Region Linguistic Distance

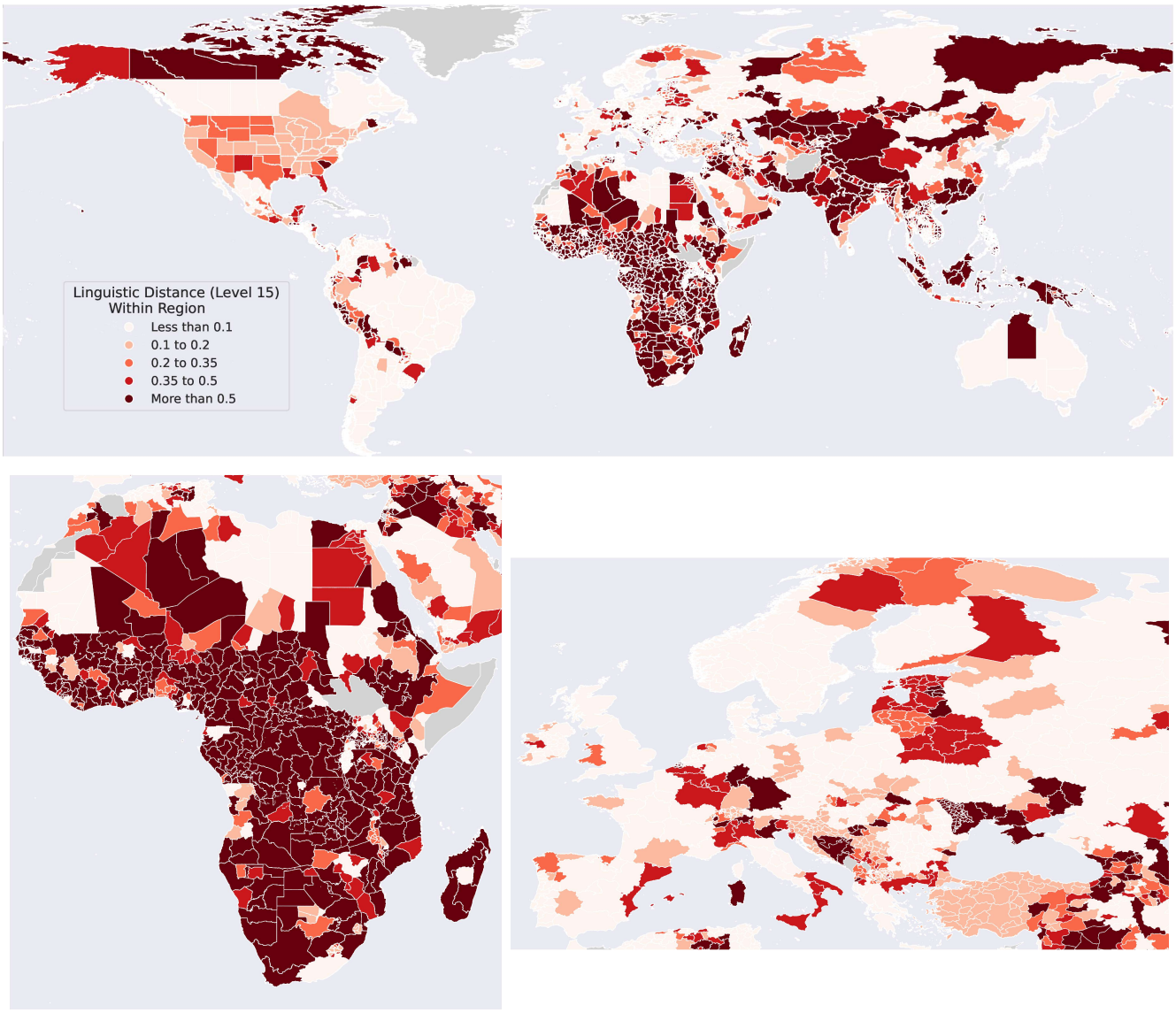


Figure depicts the linguistic distance between each subnational region and itself. It is measured as the expected distance between two randomly drawn individuals of the region. Darker colors indicate a greater within-region distance, and hence a weaker propensity to secede. Languages are defined at its finest level (aggregation level 15).

in recent decades they have only rarely translated into subnational regions becoming effectively independent. Since the breakup of the Soviet Union and Yugoslavia, the only successful secessions have been East Timor from Indonesia, Eritrea from Ethiopia, and South Sudan from Sudan. Not only have we seen limited changes to the world map due to secessions of subnational regions, we have also witnessed few modifications due to unions between countries. In the last thirty years, the only internationally recognized unions have been between South and North Yemen and between East and West Germany.

Based on the observation of a relatively stable world map, we conduct a grid search over the parameter space $\{\alpha, \beta, \delta, \eta\}$ and identify the combination of parameter values that minimizes the average of the model-

Figure 3: Income per Capita Relative to Current Country

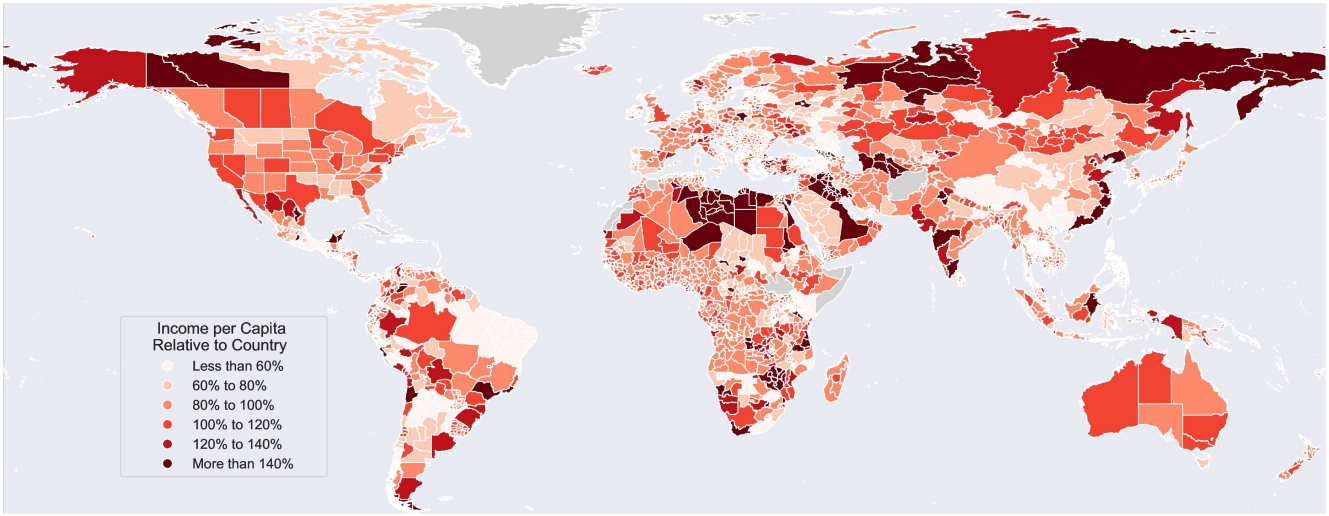


Figure depicts the income per capita of each subnational region relative to that of the country it belongs to in the year 2000. Darker colors indicate a relatively richer subnational region, and hence a stronger propensity to secede.

predicted share of potential secessions and the model-predicted share of potential unions. More specifically, for each combination $(\alpha, \beta, \delta, \eta)$, we determine the share of subnational regions with a majority in favor of seceding, and the share of neighboring pairs of countries with a majority in both neighbors in favor of uniting. For the grid search we consider the parameter space $\beta \in [0.05, 0.95]$ in increments of 0.001, $\delta \in [0.05, 0.95]$ in increments of 0.001, and $\eta \in (1, \ldots, 15)$. Recall that the value of α is immaterial to the support for secessionism, so we set its value to 1. This procedure yields parameter values $\beta = 0.182$, $\delta = 0.198$ and $\eta = 15$.

Alternative calibration. In an alternative calibration, we target the values of α and β to data on defense spending, and we target the values of δ and η to the regionalist vote share in a subgroup of advanced democracies. More specifically, for any given combination of δ and η, we choose the values of α and β that minimize the sum of squared errors between model-predicted government spending and observed average defense spending in developed countries between 1995 and 2015 according to the *Government Finance Statistics Yearbook* of the International Monetary Fund. We then choose the values of δ and η that best predict the secessionist vote share in the three countries with the highest regionalist vote share between 1995 and 2014 according to Sorens (2008). This calibration yields parameter values $\alpha = 8.0902$, $\beta = 0.1264$, $\delta = 0.16$ and $\eta = 9$.

When targeting public spending, we focus on defense, because it comes close to what might be considered a country-level public good. One possible concern is that secessionist conflict may drive defense spending. To avoid this concern, we only consider developed countries with a GDP per capita of at least 50% of that of the European Union. Another possible concern is that we are not giving a big enough role to public goods, because defense spending typically does not make up more than 1-2% of GDP. However, if we were to multiply defense spending by, say, a factor of 10, this would simply translate into a different calibrated α. As we have shown in (11), this would leave the population share in favor of secession unchanged. One might also be concerned that calibrating to developed countries in the context of quasi-linear preferences implies excessive predicted public spending in developing countries. While this could easily be addressed by setting α to a lower value in less developed countries,

there is no need to do so, since the results are invariant to changes in α.

4 Predicted Propensity to Secede

This section explores the predictions of our calibrated model for the instability of countries.

Baseline calibration. Using our baseline calibration, Figure 4 shows the model-based share of each region's population in favor of secession. The top 10% of regions with strongest support for secessionism include regions such as Tibet (China), Southern Nations (Ethiopia), Bavaria and Saarland (Germany), Aceh (Indonesia), Lombardia and Sardinia (Italy), Okinawa (Japan), Friesland (Netherlands), Arad (Romania), Tatarstan (Russia), Western Cape (South Africa), and Galicia and Catalonia (Spain). Some of these regions have witnessed violent territorial conflict (Aceh and Tibet), some feature significant pro-independence movements (Catalonia and Okinawa), some have distinct regional cultures (Friesland, Galicia and Sardinia), and others are border regions that in the past were part of a neighboring country (Saarland).

For the same baseline calibration, Figure 5 depicts the model-based share of each country's population that prefers their subnational region to secede. Ignoring small island nations, the most unstable countries for each of the six major continents are India, Italy, Papua New Guinea, South Sudan, Guatemala, and Peru. When comparing secessionism by the world's main regions as classified by the World Bank, the highest support in terms of population share is found in South Asia (14%), followed by sub-Saharan Africa (13%). The lowest support is found in North America, where less than 1% of the population favors secession. Appendix Table B.1 gives the full list of countries, and provides additional measures of instability, such as the share of regions with a majority in favor of secession and the share of the population residing in regions with a majority in favor of secession.

Alternative calibration. Turning to the alternative calibration, Appendix Figure B.3 displays the regional and country population shares in favor of secession, and Appendix Table B.2 gives the full list of countries with their respective instability measures. The most unstable countries for each of the six continents are Pakistan, Belgium, Papua New Guinea, Ethiopia, Canada and Bolivia. The correlation between the country population shares in favor of secession in the baseline calibration and the alternative calibration is 0.59. One difference is that there is a slightly smaller average country population share in favor of secession (6.2%, compared to 6.9% in the baseline calibration).[12] One reason is that linguistic groups are measured at a coarser level (η is 9, compared to 15 in the baseline).

5 Assessing Model Performance

Before conducting counterfactual exercises to identify the relative importance of identity and income in driving secessionism, we analyze the performance of our calibrated model. We do so by correlating our model-predicted measures of instability to actual measures of instability and by exploring whether the calibrated model can account for the breakup of the Soviet Union.

[12] This may come as a surprise, since the baseline calibration targets the stability of the current world map. However, in that baseline calibration, we minimize not just the share of regions with a majority in favor of secession, but also the share of neighboring country pairs with majorities in favor of uniting.

Figure 4: Model-Based Share of Regional Population in Favor of Secession

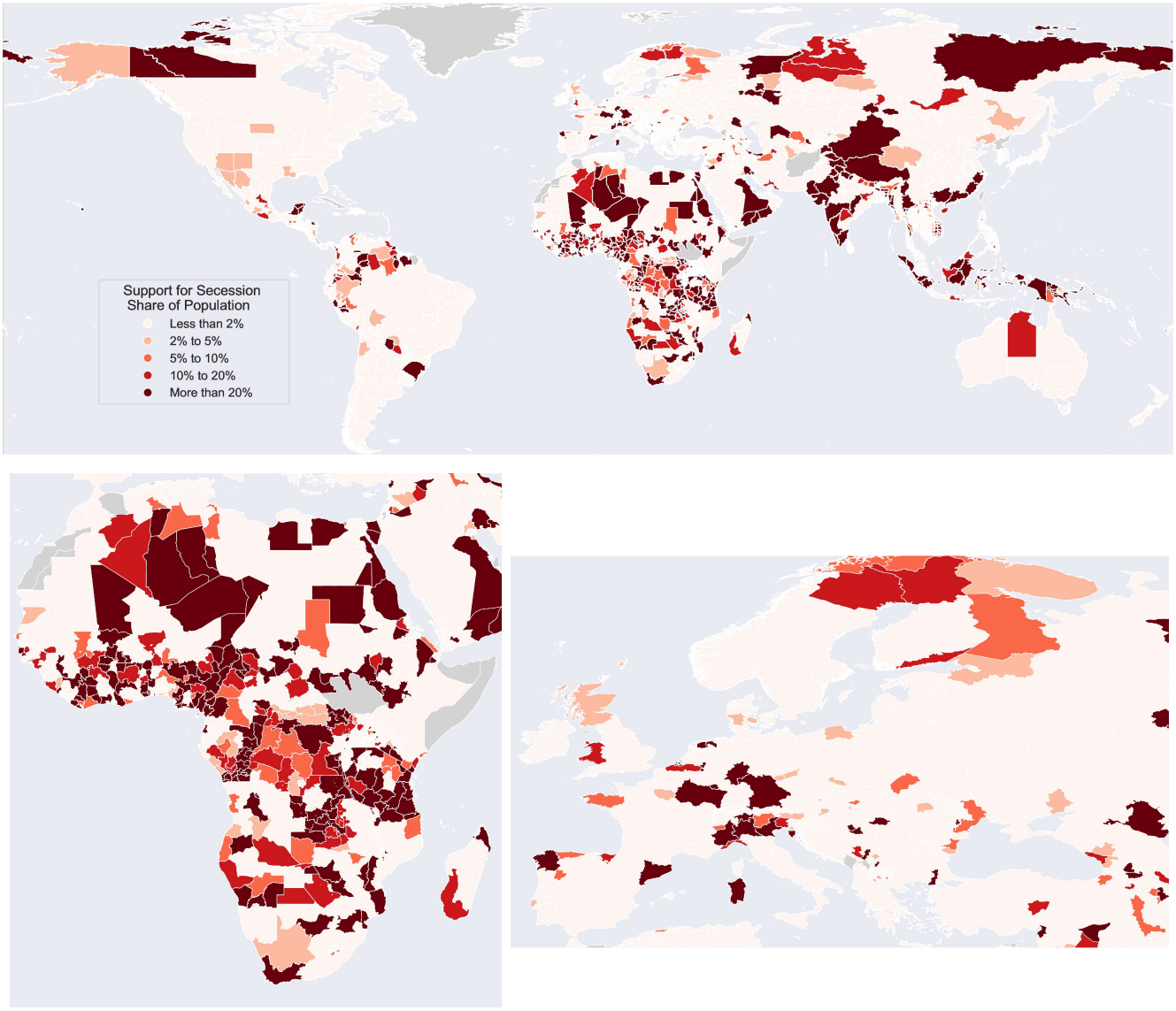


Figure depicts the model-generated share of the population in favor of secession for each of the 3,003 subnational regions. Results are based on the baseline calibration.

5.1 Correlation with Actual Measures of Instability

This section compares our model-based measures of instability to actual secessionist activity, state fragility, regional autonomy, and conflict.

Secessionist movements. The most direct way of assessing the accuracy of our quantitative model is to compare the propensity to secede as predicted by the model to actual secessionist activity around the world. To that end, we compile a novel global dataset on the presence, number, intensity, and geographic distribution of secessionist movements. Our starting point is a global list of 2,529 active secessionist movements recorded by

Figure 5: Model-Based Share of Country Population in Favor of Secession

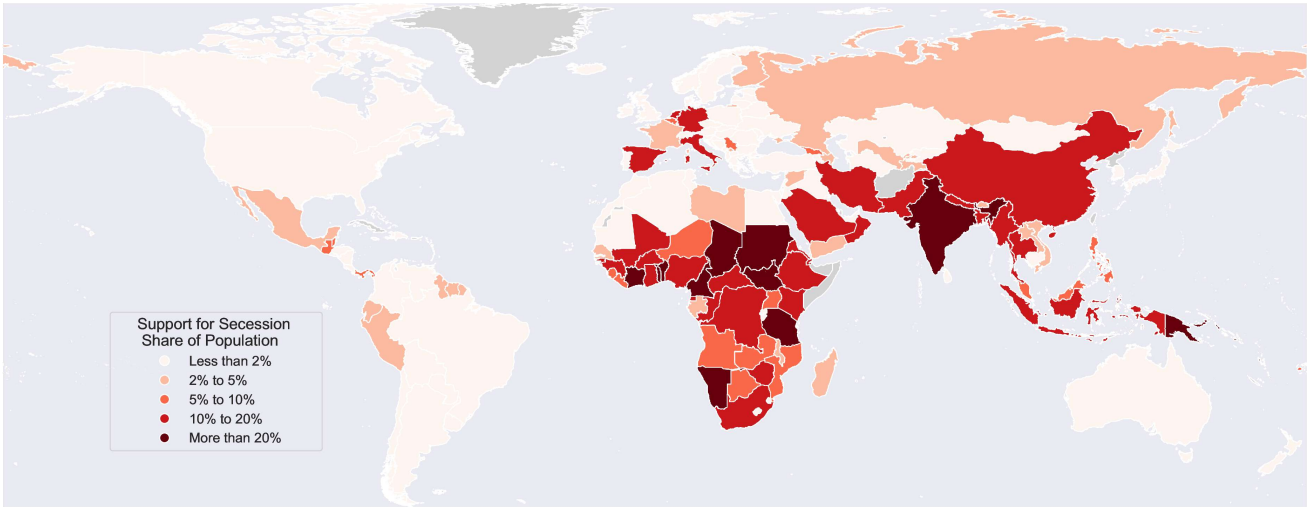


Figure depicts the model-generated share of the population in favor of secession for each country. Results are based on the baseline calibration.

Wikipedia.[13] These secessionist movements include political parties, militant and civil organizations, and social and ethnic movements. Some are armed and violent; others are non-violent and unarmed. As an example, for the case of Catalonia in Spain, the list includes political parties, such as Esquerra Republicana de Catalunya and Junts per Catalunya, civil organizations, such as Assemblea Nacional Catalana and Òmnium Cultural, and youth organization, such as Arran and La Forja. As another example, for the case of Balochistan in Pakistan, it includes political parties, such as the Baloch National Movement, and militant organizations, such as the Baloch Liberation Army.

Using this global list, we link each secessionist movement to one or more first-level administrative regions. This determines for each subnational region in our dataset whether there exists any active secessionist movement. We find that secessionist movements are active in 511 first-level administrative regions in 94 countries. To get a measure of the intensity of secessionist activity by region, we use two indicators. The first is simply the number of secessionist movements by region. The second is the number of visits to the Wikipedia webpages of these secessionist movements. For this, we sum the page views over all of Wikipedia's language versions for the period 2015-2020. Using the example of Catalonia, the political party Esquerra Republicana de Catalunya received 2,037,951 views between 2015 and 2020, whereas the civil organization Assemblea Nacional Catalana had 1,212,116 visits and La Forja got only 22,838 page views.

We now explore how predictive our model-based measures of instability are of actual secessionist activity around the world. Panel A of Table 1 reports our findings for country-level regressions of actual secessionist activity on model-based measures of secessionism, controlling for World Bank region fixed effects and the number of subnational regions per country. In column (1) we see that the model-based number of secessionist regions by country is strongly predictive of the actual number of secessionist regions.[14] Columns (2) and (3) show that the

[13]The list of active secessionist movements comes from Wikipedia and was downloaded on October 12, 2020. It is organized by continent (Africa, Asia, Europe, North America, Oceania, South America).

[14]To be a secessionist region, the model requires 50% of its population to favor independence whereas the data require there to be at least one secessionist movement.

same is true for the model-based share of the population that favors separatism and the model-based measure of at least one region being in favor of secession. These different model-based measures of secessionism are also statistically significant predictors of the actual number of secessionist movements (columns (4) to (6)) and the number of Wikipedia page views of secessionist organizations (columns (7) to (10)). The magnitudes of the effects are large: countries that according to the model have at least one secessionist region have 83% more secessionist movements in the data and 285% more page views of the Wikipedia entries of these movements.

Panel B of the same table reports our findings for subnational regressions of actual secessionist activity on model-based measures of secessionism. Column (1) shows that the model-based share of the region's population in favor of secession is highly predictive of the region having at least one secessionist movement in reality. Columns (2) and (3) show that the same holds for model-based indicator variables that measure whether a region has at least 10% or 50% of its population that favors independence. These different model-based measures of secessionism are also highly predictive of intensive measures of secessionist activity across regions, such as the number of secessionist groups (columns (4) to (6)) and the number of Wikipedia page views of secessionist movements in the region (columns (7) to (10)). The magnitudes of the effects are substantial: having at least 10% of a subnational region's population in favor of secession according to the model is associated with a 18% increase in the number of secessionist groups in that region and 99% more visits to the Wikipedia entries of those groups. These results control for country fixed effects, and hence account for any unobserved heterogeneity at the country level, such as national institutions and culture. Overall, these large, positive, and statistically significant associations suggest that our model-based measures of potential demand for secession at the country and regional levels capture some of the forces underlying actual secessionist movements around the world.

As an alternative to our data on active secessionist movements, we use the database on self-determination movements by Sambanis et al. (2018), covering 120 countries for the period 1945-2012. Their data expand on previous data collection efforts on self-determination movements, such as the Minorities at Risk (MAR) project (Gurr 1993, 2000) and the Peace and Conflict reports from the Center of International Development and Conflict Management (Marshall and Gurr, 2003, 2005). Compared to our dataset, Sambanis et al. (2018) include certain ethnic groups that do not seek secession but rather autonomy or indigenous land rights (e.g., the Cacarica or Paez communities in Colombia, or the Hazaras in Afghanistan). A disadvantage is that their data are not linked to specific subnational regions, and they do not include a measure of the intensity of self-determination movements. In Appendix Table B.3 Panel A we analyze the association between our model-based measures of instability and the presence and intensity of self-determination movements during the 2000-2012 period according to Sambanis et. al (2018). The results show that our model-based measures are strong predictors of the presence of self-determination movements (columns 1-3) as well as their number (columns 4-6), after accounting for World Bank region fixed effects and the number of regions in a country. Panel B show results when we combine our data from Wikipedia with those of Sambanis et al. (2018). Specifically, we code a country as having separatist movements if it has at least one according to one of these sources. We also use the maximum number of movements across both data sources as a proxy for the actual number of movements in the country. The results are unchanged.

State fragility. Another expression of country instability is its vulnerability to collapse. The Fragile States Index (FSI), developed by the Fund for Peace, uses a conflict assessment framework to evaluate this type of

Table 1: Predicted Demand for Secession and Contemporary Secessionist Activity

Panel A: Country-Level Analysis	Interest in Secession								
	Secessionist Activity in Country						Log[Number of Wikipedia Page Views]		
	Log[1+# Secessionist Regions]			Log[1+# Secessionist Groups]			Secessionist Organization		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Log[1 + # Secessionist Regions]	0.347***			0.585***			2.144**		
	(0.113)			(0.202)			(0.894)		
Log[1 + Share Pop. Pro-Secession]		2.319***			3.966***			15.107**	
		(0.799)			(1.457)			(6.957)	
At least 1 Region Pro-Secession			0.481***			0.831***			2.847**
			(0.143)			(0.247)			(1.112)
WB Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Log[1+# Regions]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted-R^2	0.33	0.33	0.34	0.27	0.27	0.28	0.18	0.18	0.18
Observations	173	173	173	173	173	173	173	173	173

Panel B: Regional-Level Analysis	Interest in Secession								
	Secessionist Activity in Region						Log[Total Number of Wikipedia Page Views]		
	Secessionist Region			Log[1+# Secessionist Groups in Region]			Secessionist Organization		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Log[1 + Sh. Pop. Pro-Secession]	0.301***			0.454***			2.453***		
	(0.070)			(0.124)			(0.760)		
At least 10% Pro-Secession		0.119***			0.180***			0.991***	
		(0.028)			(0.049)			(0.320)	
At least 50% Pro-Secession			0.140***			0.227***			1.175***
			(0.038)			(0.080)			(0.411)
Country FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted-R^2	0.21	0.21	0.21	0.24	0.24	0.24	0.22	0.22	0.21
Observations	3003	3003	3003	3003	3003	3003	3003	3003	3003

Notes: Heteroskedasticity robust standard error estimates clustered at the country-level are reported in parentheses; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests.

vulnerability.[15] The FSI is a composite measure covering various aspects of state power and fragility.[16] Panel A of Table 2 reports the association between our model-based measures of secessionism and a country's ranking (columns (1) to (3)) and value according to the FSI (columns (4) to (6)). In all cases, the association is statistically significant at the 1% level. To give a sense of the magnitude of the effects, a country with at least one region in favor of secession changes its rank by 24 positions in the direction of greater fragility compared to one that has no secessionist region. Such a country also increases its FSI value by an average of 10 points, a large change close to half a standard deviation of FSI. Appendix Table B.4 shows that the positive association holds for all 12

[15]The data is accessible at https://fragilestatesindex.org/. Accessed on June 23, 2021.

[16]The index consists of the following subcomponents: C1: Security Apparatus, C2: Factionalized Elites, C3: Group Grievance, E1: Economic Decline, E2: Uneven Economic Development, E3: Human Flight and Brain Drain, P1: State Legitimacy, P2: Public Services, P3: Human Rights and Rule of Law, S1: Demographic Pressures, S2: Refugees and IDPs, X1: External Intervention.

subcomponents, and the association is statistically significant at the 10% for all but three subcomponents.

Table 2: Predicted Demand for Secession and State Fragility, Regional Autonomy, and Conflict

Panel A: Institutional	Fragile State Index (2006-2021)						Regional Authority Index (1950-2016)		
	Rank			Index			Total		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Log[1 + # Secessionist Regions]	19.343***			8.038***			3.784**		
	(5.346)			(2.393)			(1.567)		
Log[1 + Share Pop. Pro-Secession]		97.012***			46.765***			53.527**	
		(36.300)			(15.742)			(21.725)	
At least 1 Region Pro-Secession			24.114***			9.749***			5.208**
			(7.061)			(3.150)			(2.362)
WB Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted-R^2	0.45	0.43	0.45	0.43	0.42	0.43	0.19	0.22	0.19
Observations	167	167	167	167	167	167	89	89	89

Panel B: Conflict	Intensity of Conflict (1997-2020)								
	Log[1 + # Deaths]			Log[1 + # Events]			Log[1 + # Years]		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Log[1 + # Secessionist Regions]	1.896***			1.333***			0.296**		
	(0.404)			(0.437)			(0.116)		
Log[1 + Share Pop. Pro-Secession]		4.386			0.113			-0.060	
		(3.676)			(3.671)			(1.092)	
At least 1 Region Pro-Secession			2.268***			1.486**			0.425**
			(0.553)			(0.573)			(0.163)
WB Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted-R^2	0.45	0.38	0.44	0.24	0.19	0.22	0.66	0.64	0.66
Observations	173	173	173	173	173	173	173	173	173

Notes: Heteroskedasticity robust standard error estimates clustered at the country-level are reported in parentheses; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests.

Regional autonomy. In many countries, the demand for secessionism leads to greater regional autonomy. We therefore look at the association between our model-based measures of secessionism and the Regional Authority index (RAI), developed by Hooghe et al. (2016).[17] The index measures the authority of subnational governments in 95 democracies or quasi-democracies on an annual basis from 1950 to 2018. It captures different aspects of self-rule and shared-rule, based on ten subcomponents related to institutional depth, policy scope, fiscal autonomy, borrowing autonomy, representation, law making, executive control, fiscal control, borrowing control, and constitutional reform. The last 3 columns of Panel A in Table 2 report the association between our model-based measures of secession and the Regional Autonomy Index (RAI). Once again, the estimated associations are positive and statistically significant. For example, a country with at least one secessionist region has an RAI that is 5.2 points higher (equivalent to 0.6 of a standard deviation) compared to a country with no secessionist regions. Appendix Table B.5 shows the association between one of our model-based measures and the ten subcomponents of RAI.

[17]Data available at https://garymarks.web.unc.edu/data/regional-authority-2/. Accessed on September 16, 2021.

As can be seen, the statistically significant associations tend to be with components that are related to self-rule, rather than with shared-rule. This makes sense, given that secessionism is more concerned with self-determination than with participating in national policy.

Figure 6: Predicted Demand for Secession and Conflict Intensity (Deaths) across Countries

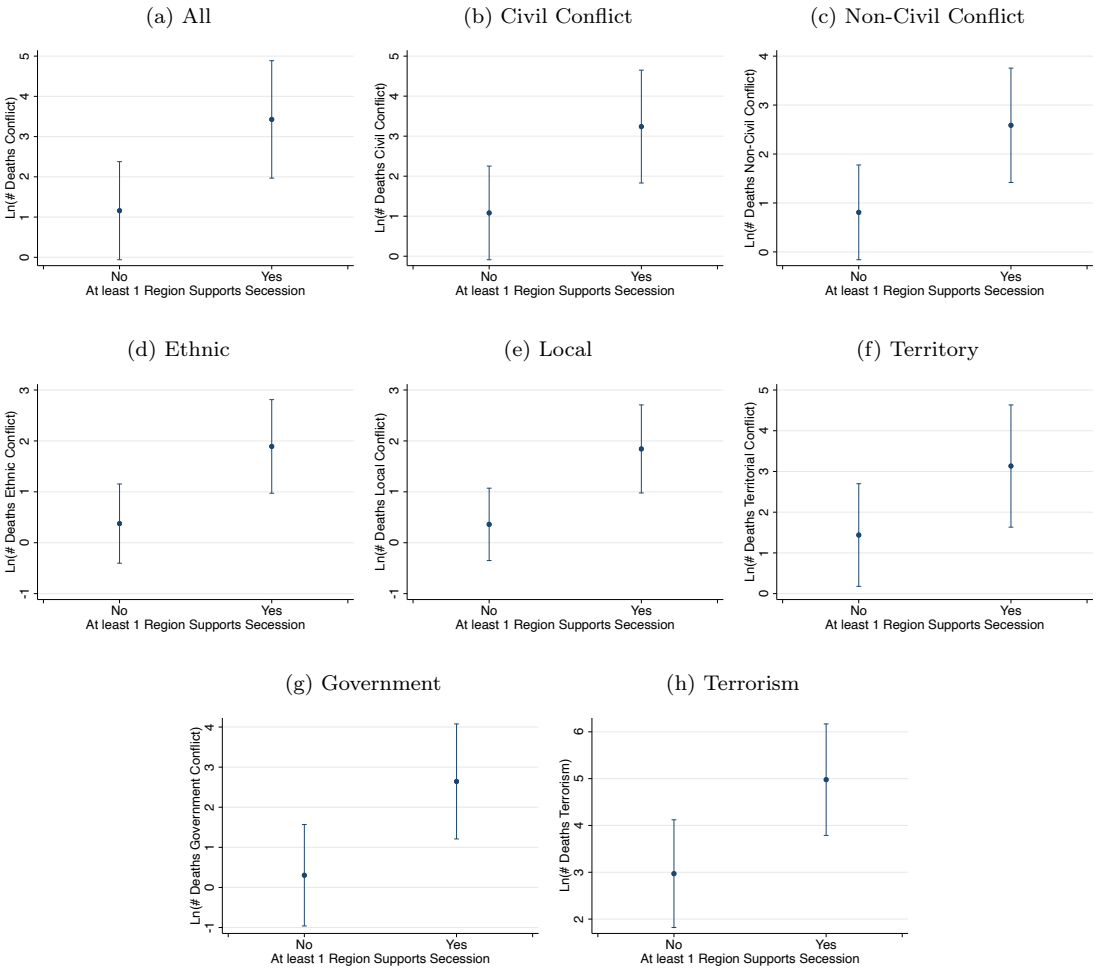


Figure shows the association between our model-based predicted demand for secession as measured by a country having at least 1 region in favor of secession and the number of deaths by type and source of conflict.

Conflict. As a further validation, we explore the association between the model-based demand for secession and conflict within countries. Our main data on conflict come from the Armed Conflict Location and Event Data-ACLED (Raleigh et al. 2010). Panel B in Table 2 shows that most of our measures are positively and significantly associated with conflict intensity within countries. For example, countries that we predict have at least one secessionist region have on average 227% more deaths, are involved in 149% more conflict events, and suffer from conflict for 43% more years. Using additional conflict data from the UCDP/PRIO Armed Conflict Dataset (Gleditsch et al., 2002)) and terrorism data from the Global Terrorism Database (GTD),[18] Figure 6 shows

[18]GTD is available at https://www.start.umd.edu/gtd/. Analysis uses the September 2019 version accessed on November 6, 2019.

that the results hold for various types of conflict, including civil, ethnic and local, as well as for terrorism events.

Alternative calibration. Appendix Tables B.6-B.10 and Appendix Figure B.4 report the same validation analysis for our alternative calibration. The results are qualitatively similar: our predicted measures of secessionism are positively and significantly associated with actual secessionist activity, state fragility, regional autonomy and conflict. Overall, from these different comparisons with the data, we can conclude that our quantitative model performs well in predicting potential secessionist activity around the globe.

5.2 The Predicted Disintegration of the Soviet Union

Between 1990 and 1991, the 15 different republics that made up the Soviet Union became independent countries. As a final validation exercise, we analyze whether our calibrated model would have predicted the disintegration of the Soviet Union. More specifically, using data from the end of the 1980s, we compute for each of the republics the share of the population in favor of secession.

It is not immediately obvious how to get reliable data on income per capita from the Soviet era. In principle, we could use the G-Econ 4.0 data, because they go back to 1990. However, for most Soviet republics, the 1990 data are based on 1995 or later, and projected backward using a variety of methods. Given this drawback, we rely on the monograph by Flakierski (1992) who reports wage data of manual and non-manual labor for each Soviet republic in 1988. Since wages are not the same as income per capita, we use data on real GDP per capita from Russia in 1990 to rescale the wage data. As long as the ratio between GDP per capita and wages does not differ much across republics, this gives us a reasonable proxy of income per capita for 1988. For population data by republic, we use the 1989 Soviet Census.

Table 3: Predicted Secession of Soviet Republics

Model-Predicted Share of Population in Favor of Secession from the Soviet Union							
Republic Wage Data 1988		Regional Wage Data 1988		GDP per Capita Data 1990		Alternative Calibration	
(1) Country	(2) Share	(3) Country	(4) Share	(5) Country	(6) Share	(7) Country	(8) Share
Uzbekistan	0.02	Uzbekistan	0.02	Armenia	0.00	Belarus	0.00
Belarus	0.25	Tajikistan	0.03	Azerbaijan	0.00	Kyrgyzstan	0.01
Kazakhstan	0.46	Belarus	0.25	Belarus	0.00	Ukraine	0.02
Kyrgyzstan	0.54	Kazakhstan	0.46	Kyrgyzstan	0.00	Kazakhstan	0.43
Latvia	0.66	Kyrgyzstan	0.54	Moldova	0.00	Latvia	0.66
Moldova	0.73	Moldova	0.64	Tajikistan	0.00	Moldova	0.73
Estonia	0.75	Latvia	0.66	Turkmenistan	0.00	Tajikistan	0.75
Tajikistan	0.75	Ukraine	0.72	Uzbekistan	0.00	Estonia	0.75
Ukraine	0.77	Estonia	0.75	Kazakhstan	0.09	Georgia	0.78
Georgia	0.78	Georgia	0.78	Ukraine	0.66	Uzbekistan	0.81
Azerbaijan	0.86	Azerbaijan	0.86	Latvia	0.66	Azerbaijan	0.86
Lithuania	0.87	Lithuania	0.87	Estonia	0.75	Lithuania	0.87
Turkmenistan	0.91	Turkmenistan	0.91	Georgia	0.76	Turkmenistan	0.91
Armenia	0.95	Armenia	0.95	Lithuania	0.87	Armenia	0.95

Notes: Columns (1) and (2) are based on republic-level wage data of 1988 from Flakierski (1992). They have been scaled by the ratio of income per capita to wages in Russia in 1990. Columns (3) and (4) are based on the same data, but allow for within-repubilc wage differences between administrative regions by using data on relative income per capita of 1990 from G-Econ 4.0. Columns (5) and (6) are based on 1990 income per capita data from G-Econ 4.0. Columns (7) and (8) are based on the same data as columns (1) and (2), but use the parameters of the alternative calibration.

Using these data, what does the calibrated model predict? Columns (1) and (2) of Table 3 show a majority in favor of secession in all but three Soviet republics: Uzbekistan, Belarus, and Kazakhstan.[19] We conduct two further robustness checks. First, our income per capita proxy is at the level of each republic, because the Flakierski (1992) wage data are at the level of republics. To generate wage differences between subregions of the republics, we set the relative wage levels of the different subregions to those of income per capita of 1990 in G-Econ 4.0. The results are largely unchanged. Columns (3) and (4) show that one additional republic, Tajikistan, now prefers to stay within the Soviet Union. Second, in spite of the drawbacks, we directly use the income per capita data from G-Econ 4.0. In that case, five more republics prefer to remain within the Soviet Union: Armenia, Azerbaijan, Kyrgyzstan, Moldova and Turkmenistan. In all three exercises, the three Baltic republics (Lithuania, Letonia and Estonia), Georgia and Ukraine prefer to leave the Soviet Union. And in all three exercises, Uzbekistan, Belarus, and Kazakhstan prefer to remain within the union.

How do these results compare to the historical record? One basic finding is that our model predicts that the Soviet Union was clearly unstable. Given that in our quantitative model only 5.9% of the 3,003 subnational regions have a majority in favor of secession, this was not a foregone conclusion.

To further investigate how well the model fits the historical record, we also compare our findings to the timeline of the breakup of the Soviet Union and to the emergence of alliances in the post-Soviet world. In terms of timeline, in March 1990, Lithuania became the first republic to declare independence from the Soviet Union. The other two Baltic states, Latvia and Estonia, followed suit in the subsequent months. The first non-Baltic republic to secede was Georgia in April 1991. In the ensuing Summer and Fall, most remaining republics became independent, with only Belarus, Russia and Kazakhstan remaining in the union. Those last three finally also became independent countries in December 1991. Our findings are remarkably consistent with this timeline: the first four republics to declare independence have a majority in favor of secession in all of our exercises, and the two republics to remain in the union until the very end fail to reach a majority in all three exercises.

In the post-Soviet era, many cooperation agreements surfaced between subsets of the successor states. Perhaps the two most important ones are the Commonwealth of Independent States (CIS) and the Eurasian Economic Union (EAEU). The CIS is an intergovernmental organization that promotes economic, political, and military cooperation between its member states. Its current members are Armenia, Azerbaijan, Belarus, Kazakhstan, Kyrgyzstan, Moldova, Russia, Tajikistan and Uzbekistan. The EAEU is an economic union between Armenia, Belarus, Kazakhstan, Kyrgyzstan and Russia. Consistent with what our findings would suggest in all three exercises, Georgia and Ukraine do not participate in either of these organizations, whereas Kazakhstan and Belarus are members of both.

As a last robustness check, we redo our first exercise using the parameters of the alternative calibration. In this case, the model predicts that all but four republics want to secede. As before, the first four republics to declare independence (Latvia, Lithuania, Estonia and Georgia) have a majority in favor of independence, and the last two republics to remain in the union (Belarus and Kazakhstan) have less than 50% supporting secession. The only significant difference with our other exercises is that Ukraine now chooses to remain inside the Soviet Union. The reason for this difference is straightforward: under the alternative calibration, Ukrainian and Russian are aggregated into the same linguistic identity group, weakening the incentive to secede.

[19] We did not assess the propensity of Russia to secede. As the central player of the Soviet Union, its incentives to keep the union together were different from those of the other republics.

6 Is Secessionism Mostly about Identity or Income?

In the model, both identity and income per capita determine the incentives to secede. In this subsection we analyze their relative importance.

Identity and income per capita differences. We start by conducting two counterfactual exercises. In a first exercise, we remove differences in subnational income per capita by assuming that each region's income per capita is equal to the country average. In that case, the incentives to separate depend on identity. In a second exercise, we remove identity differences by assuming that everyone speaks the same language.[20] In that case, the incentives to secede depends on income per capita. Appendix Table B.11 reports the results of these counterfactual exercises. More specifically, it gives the change in a country's population share in favor of secession when either only identity matters or only income per capita matters.

When removing differences in income per capita and maintaining identity as the main determinant, the average support for secession does not drop. In fact, it even increases, by 0.6 percentage points at the country level (from a baseline of 6.9%) and by 0.9 percentage point at the regional level (from a baseline of 7.5%). This may come as a surprise, since secessionism tends to be stronger in regions with higher income per capita. Hence, we would expect that equalizing income per capita should on average weaken the propensity to secede. However, there are also many subnational regions with lower income per capita and a distinct identity. In those regions, equalizing income per capita tends to strengthen the support for independence. Quebec provides a good example. There, setting income per capita to the Canadian average leads to an increase in secessionist sentiment. It turns out that examples such as Quebec dominate, so that eliminating within-country income per capita differences actually slightly increases the average support for secession.

In contrast, when removing identity differences and maintaining differences in income per capita, the support for secession drops on average by 6.4 percentage points at the country level (from a baseline of 6.9%) and by 6.9 percentage points at the regional level (from a baseline of 7.5%). This implies that in many countries secessionism loses all its support. Again, there are a few exceptions: in Namibia 14.5% of the population continues to support secession (down from 20.2% in the baseline) and in Sudan support is at 17.2% (down from 22.4% in the baseline).

Taken together, these two counterfactual exercises suggest that identity is much more important than income per capita for secessionism. One potential issue with this conclusion is that the unimportance of income per capita might be due to averaging across subnational regions. For example, when we equalize income across regions in South Africa, secessionism collapses in the relatively rich Western Cape, whereas it emerges in the relatively poor Eastern Cape. To investigate whether the absence of income per capita as a significant driver of secessionism is due to averaging, we limit our attention to the subgroup of regions with a strictly positive share of the population in favor of secession in the baseline. When removing income per capita differences, the support of secessionism in this subgroup of regions does drop, but not by much, from a baseline average of 44% of the population to 38%. This confirms that secessionism is not very sensitive to changes in income per capita.

These findings are confirmed when using the alternative calibration. In that case, when removing within-country differences in income per capita, the average support for secession hardly changes, from 6.4% of a region's population

[20] We consider two different methods of eliminating identity differences in the context of our model. A first sets δ to 0, and hence $S(\ell, C)^{\delta}$ to 1, in equation (1). In that case, the contribution of the public good to the utility of an individual of group ℓ residing in region r of country C is $\alpha G(C)^{\beta}$. A second sets $S(\ell, C)^{\delta}$ to its population-weighted country average $\sum_{\ell} S(\ell, C)^{1+\delta}$. In that case, the contribution of the public good to the utility of an individual of group ℓ residing in region r of country C is $\alpha \left(\sum_{\ell} S(\ell, C)^{1+\delta}\right) G(C)^{\beta}$. With both methods, the utility of the public good no longer depends on the linguistic group one belongs to. Below we report results based on the first method, but the results are qualitatively unchanged when using the second method.

to 6.3%. In contrast, when removing within-country identity differences, the average support for secession collapses to 0.5%. Once again, we might worry that the weak role of income per capita is due to averaging across subnational regions. This is not the case: when only considering the subgroup of regions with strictly positive support for secession in the baseline, we find that secessionism falls only slightly, from an average of 47% of a region's population to 41%.

All of this suggests that identity is the essential driver of secessionism, with economic forces playing a much smaller role. Next, we conduct further counterfactual analysis to deepen our understanding of the relative importance of the different drivers of secessionism.

Sensitivity of secessionism to changes in income per capita. In the baseline calibration, subnational regions with at least 10% of their population in favor of secession are on average 13% richer than the countries they belong to. Our counterfactual exercises so far have shown that removing this income advantage does not suffice for secessionism to weaken. Might this be because setting their income per capita to the national average is only a small shock? How far should income per capita fall for those subnational regions to give up on secessionism?

To answer this question, we look at each subnational region with at least 10% of its population in favor of secession, and compute by how much income per capita needs to decrease for secessionist support to fall below 1%.[21] The maps in Figure 7 depict our findings. In Catalonia (Spain), a drop in income per capita by 5% is enough for secessionism to vanish. For most of the other examples we discussed before, the corresponding figures tend to be larger: Tibet (China), -30%; Southern Nations (Ethiopia), -40%; Saarland (Germany), -35%; Aceh (Indonesia), -35%; Lombardia (Italy), -35%; Sardinia (Italy), -40%; Okinawa (Japan), -30%; Tatarstan (Russia), -30%; Western Cape (South Africa), -20%.

Going beyond specific examples, income per capita needs to drop on average 43% for secessionist support to vanish in subnational regions with at least 10% of its population in favor of secession. Using the alternative calibration, the corresponding drop in income per capita is 44%. Where does this leave us in terms of assessing the importance of economic forces? On the one hand, income per capita matters: poor regions have less incentives to separate. On the other hand, subnational regions need to be economically far behind for secessionism to substantially weaken. Therefore, secessionism is only very weakly sensitive to changes in income per capita. This confirms our previous conclusion.

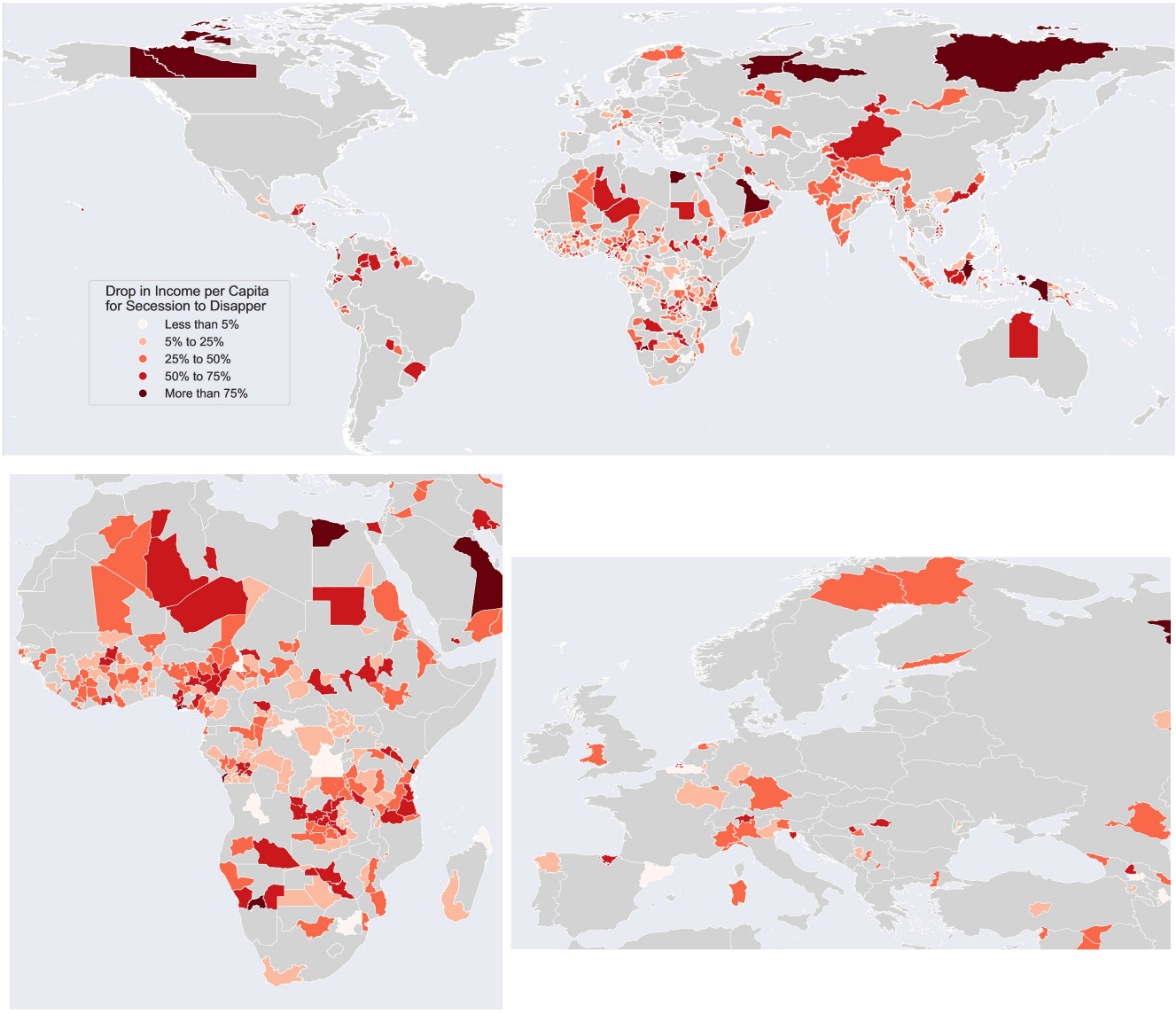
Sensitivity of secessionism to within-region identity heterogeneity. While identity differences with the rest of the nation strengthen secessionist tendencies, identity differences within subnational regions diminish secessionist tendencies. Indeed, regions have less reason to become independent if their own identity is diverse. To further assess the importance of identity for secessionism, we focus on this within-region identity heterogeneity. More specifically, we explore what happens if individuals ignore within-region identity differences in the case of independence. In this counterfactual exercise, we do not change the linguistic composition of subnational regions. Instead, we simply assume that individuals cease to care about within-region identity differences in the case of secession.[22] This amounts to setting δ equal to zero if a region chooses to become independent, while maintaining the original parameter value if a region remains part of the union.

The maps in Figure 8 depict the percentage point increase in support for secessionism if individuals ignore

[21] Of course, when a region's income per capita declines, it also shrinks the country's overall economy.

[22] An alternative exercise would be to have everyone in a region adopt a common language. However, in that case the results would depend on which language becomes the common language.

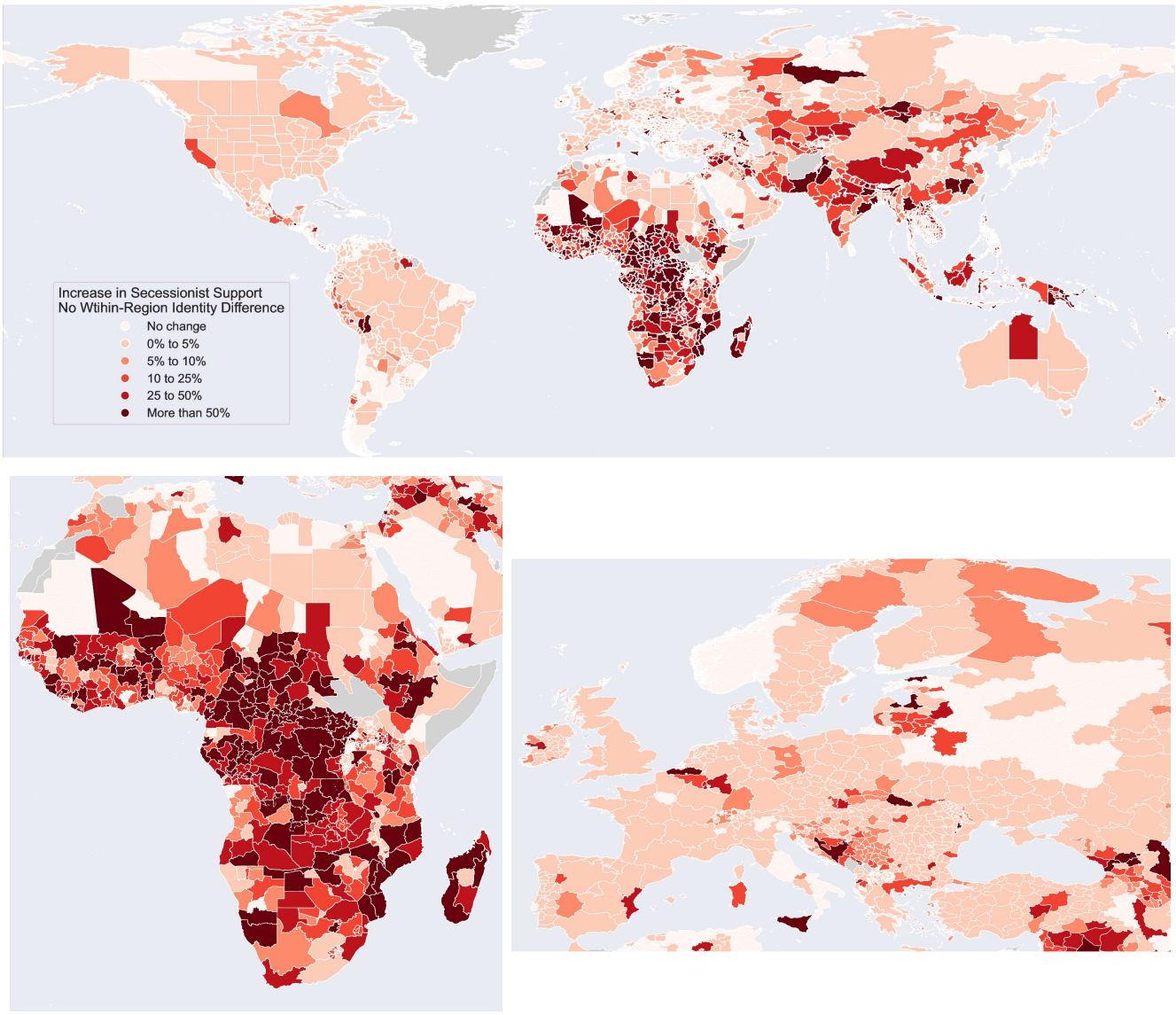
Figure 7: Drop in Income per Capita for Secessionism to Disappear



For each subnational region with more than 10% of its population in favor of secession, Figure depicts the necessary drop in income per capita for the support of secession to drop below 1%.

within-region identity heterogeneity in the case of independence. We observe a large increase in support for secessionism. At the country level, we witness a more than tripling of the support for regional separation: from 6.9% in the baseline to 24.0% in this counterfactual exercise. And at the regional level, the support increases from 7.5% to 20.0%. Under the alternative calibration, the relative increase at the country level is similar: from 6.2% in the baseline to 18.2%. The high degree of sensitivity of secessionism to within-region identity differences confirms our previous conclusion: identity is a key driver of separatist sentiment.

Figure 8: Increase in Secessionist Support when Eliminating Within-Region Identity Differences



For each subnational region, Figure depicts the increase in support for secession if people were to ignore within-region identity differences in the case of independence.

Secessionism and policy. Our counterfactual exercises strongly suggest that identity, more than economics, is key to understanding secessionist tendencies. Although economic policy might be of some help to stave off secessionist threats, creating a common identity is more likely to lead to territorial stability. This is reminiscent of the nation-building efforts of the 19th century. Alesina, Giuliano and Reich (2021) describe how the introduction of a "national language", often through compulsory education, is central to nation-building efforts. For example, at the time of Italian unification, at most 10% of its population spoke Italian. Linguistic homogenization was seen as key to keep the newly-created country united. Likewise, the expansion of the Russian Empire during the late-nineteenth century was accompanied by Russification through education. France went through a similar process

of nation-building. As Weber (1979) argues, the village school was the "ultimate acculturation process that made the French people French". Of course, the relevant question is to what extent the homogenization of language use weakens secessionist tendencies. Recent work by Blanc and Kubo (2021) show that French municipalities that benefited more from state-sponsored education in the nineteenth century displayed greater participation in the Resistance during WWII and received more votes against the 1969 Referendum on Regionalization. Needless to say, public education may also instill a sense of nationhood and a common identity besides the adoption of a common language.

7 Conclusions

This paper analyzed whether the demand for secession is mostly driven by economic differences or identity differences. To study this question, we proposed a model where the tax rate is determined by majority vote and where the utility an individual derives from the public good is declining in how distinct her identity is from the rest of the country. In equilibrium, subnational regions that are either richer or more distinct in terms of identity pay a higher tax rate than its residents would like.

Taking the model to the data for 3,003 subnational regions covering the globe, we generated measures of the instability of countries and the propensity to secede of subnational regions. To validate our model's predictions, we constructed a novel worldwide database of active secessionist movements at the subnational level, and found that our model-based measures of instability are strongly associated with observed secessionist activity. Lending further credibility to the model, we also showed a strong association of our measures with the fragility of states, violent conflict, and regional autonomy. In addition, the model is successful at predicting the breakup of the Soviet Union.

We used counterfactual analysis to gauge the relative importance of income per capita and identity in driving the demand for independence. We did so in three ways. First, we compared what happens to the propensity to secede when removing either income per capita or identity differences, and found that the demand for independence is mostly driven by identity differences. Second, we evaluated by how much income per capita would have to drop for secessionist regions to give up on their demands. We found that large drops in income are needed, suggesting that economic differences only matter if they are large. Third, we assessed the effect of removing within-region identity differences, finding an important increase in the drive for independence. Overall, these different exercises strongly suggest that identity trumps income in determining secessionist tendencies.

While language is arguably the most important identity marker of population groups, it is not the only one. Ethnicity and religion are also key dimensions of identity. To the extent that these alternative identity markers coincide with language, as is the case for many ethnic groups in Africa and Asia, our data are sufficient. But if these cleavages cross-cut, then more data are needed to get a comprehensive picture of identity groups around the world. At this point, no good data exist that capture the multi-dimensionality of identity groups at a subnational level for the entire globe. Collecting such data should be the subject of future research.

In our work we have focused on the propensity to secede by first-level administrative regions. A more complex analysis might also consider the possible secession of coalitions of first-level administrative regions or of subdivisions of those regions. Moreover, in some cases secession might be driven by the desire to join another country, rather than by becoming independent. Our framework can be used to study these questions as well.

References

[1] Akerlof, G. A. and E. Kranton, 2000. "Economics and Identity," *Quarterly Journal of Economics*, 15(3): 715-753.

[2] Alesina, A., Giuliano, P. and B. Reich, 2021. "Nation-Building and Education," *Economic Journal*, 131, 2273-2303.

[3] Alesina, A. and E. Spolaore, 1997. "On the Number and Size of Nations," *Quarterly Journal of Economics*, 112(4): 1027-1056.

[4] Alesina, A. and E. Spolaore, 2003. *The Size of Nations*, Cambridge, MA: MIT Press.

[5] Alesina, A., Spolaore, E. and R. Wacziarg, 2000. "Economic Integration and Political Disintegration," *American Economic Review*, 90(5): 1276-1296.

[6] Álvarez Pereira, B., Portos, M. and J. Vourdas, 2018. "Waving Goodbye? The Determinants of Autonomism and Secessionism in Western Europe," *Regional Studies*, 52(2): 197-211.

[7] Blanc, G. and Kubo, M., 2021. "Schools, Language, and Nations: Evidence from a Natural Experiment in France," Brown University.

[8] Bolton, P. and G. Roland, 1997. "The Breakup of Nations: A Political Economy Analysis," *Quarterly Journal of Economics*, 112(4): 1057-1090.

[9] Bonomi, G., Gennaioli, N. and G. Tabellini, 2021. "Identity, Beliefs, and Political Conflict," *Quarterly Journal of Economics*, 136(4): 2371-2411.

[10] Clots-Figueras, I. and P. Masella, 2013. "Education, Language and Identity," *Economic Journal*, 123(570): 332-357.

[11] Comerford, D. and J. V. Rodríguez Mora, 2019. "The Gains from Economic Integration," *Economic Policy*, 34(98): 325-360.

[12] Desmet, K., Gomes, J. and I. Ortuño Ortín, 2020. "The Geography of Linguistic Diversity and the Provision of Public Goods," *Journal of Development Economics*.

[13] Desmet, K. Le Breton, M., Ortuño Ortín, I. and S. Weber, 2011. "The Stability and Breakup of Nations: A Quantitative Analysis," *Journal of Economic Growth*, 16, 183-213.

[14] Desmet, K., Ortuño Ortín, I. and R. Wacziarg, 2012. "The Political Economy of Linguistic Cleavages," *Journal of Development Economics*, 97, 322-338.

[15] Desmet, K. and Wacziarg, R., 2021. "The Cultural Divide," *Economic Journal*, 131(637): 2058-2088.

[16] Dhingra, S., Huang, H., Ottaviano, G., Pessoa, J. P., Sampson, T. and J. Van Reenen, 2017. "The Costs and Benefits of Leaving the EU: Trade Effects," *Economic Policy*, 32(92): 651-705.

[17] Fearon, J. D., 2003. "Ethnic and Cultural Diversity by Country," *Journal of Economic Growth*, 8: 195-222.

[18] Flakierski, H., 1992. "Income Inequalities in the Former Soviet Union and Its Republics," *International Journal of Sociology*, 22(3): 1-87.

[19] Fukuyama, F., 2018. *Identity : The Demand for Dignity and the Politics of Resentment*, New York: Farrar, Straus and Giroux.

[20] Gancia, G., Ponzetto, G. and J. Ventura, 2021. "Globalization and Political Structure," UPF.

[21] Gleditsch, N. P., Wallensteen, P., Eriksson, M., Sollenberg, M. and H. Strand, 2002. 'Armed Conflict 1946-2001: A New Dataset," *Journal of Peace Research*, 39(5): 615-637.

[22] Gershman, B. and D. Rivera, 2018. "Subnational Diversity in Sub-Saharan Africa: Insights from a New Dataset," *Journal of Development Economics*, 133: 231-263,

[23] Gethin, A., Martínez-Toledano, C. and T. Piketty, 2021. "Brahmin Left versus Merchant Right: Changing Political Cleavages in 21 Western Democracies, 1948-2020," unpublished manuscript.

[24] Gurevich, T., Herman, P. R., Toubal, F. and Y. V. Yotov, 2020. "One Nation, One Language? Language Diversity, Trade and Welfare," unpublished manuscript.

[25] Gurr, T. R., 1993. *Minorities at Risk: A Global View of Ethnopolitical Conflicts*, Washington, DC: US Institute of Peace Press.

[26] Gurr, T. R., 2000. *Peoples versus States: Minorities at Risk in the New Century*, Washington, DC: US Institute of Peace Press.

[27] Hooghe, L., Marks, G., Schakel, A. H., Chapman-Osterkatz, S., Niedzwiecki, S. and S. Shair-Rosenfield, 2016. *Measuring Regional Authority. Volume I: A Postfunctionalist Theory of Governance,* Oxford: Oxford University Press.

[28] Johnson, N. D., 2019. "Taxes, National Identity, and Nation Building: Evidence from France," unpublished manuscript.

[29] Li, S. X., 2010. "Social Identities, Ethnic Diversity, and Tax Morale," *Public Finance Review*, 38(2): 146-177.

[30] Marshall, M. G. and T. R. Gurr, 2003. *Peace and Conflict 2003: A Global Survey of Armed Conflicts, Self-determination Movements, and Democracy*, College Park, MD: Center for International Development and Conflict Management.

[31] Marshall, M. G. and T. R. Gurr, 2005. *Peace and Conflict 2005: A Global Survey of Armed Conflicts, Self-determination Movements, and Democracy*, College Park, MD: Center for International Development and Conflict Management.

[32] Raleigh, C., Linke, A., Hegre, H. and J. Karlsen, 2010. "Introducing ACLED-Armed Conflict Location and Event Data," *Journal of Peace Research*, 47(5): 651-660.

[33] Sambanis, N., Germann, M. and A. Schädel, 2018. "SDM: A New Data Set on Self-Determination Movements with an Application to the Reputational Theory of Conflict," *Journal of Conflict Resolution*, 62(3): 656-686.

[34] Sambanis, N. and B. Milanovic, 2014. "Explaining Regional Autonomy Differences in Decentralized Countries," *Comparative Political Studies*, 47(13): 1830-1855.

[35] Sorens, J., 2008. "Regionalists against Secession: The Political Economy of Territory in Advanced Democracies," *Nationalism and Ethnic Politics*, 14, 325-360.

[36] Suesse, M., 2019. "Adjusting the Size of Nations: Empirical Determinants of Separatism and the Soviet Breakup," *Journal of Comparative Economics*, 47(1): 50-64.

[37] Vanschoonbeek, J., 2020. "Regional (In)stability in Europe: A Quantitative Model of State fragmentation," *Journal of Comparative Economics*, 48(3): 605-641.

[38] Weber, E., 1979. *Peasants into Frenchmen: Modernization of Rural France, 1870- 1914,* Chatto and Windus: London.

Appendix

A Data Sources

First-level administrative regions. Global geographic distribution of administrative areas of all countries, at first levels of sub-division. Source: The Global Administrative Division Mapping project - GADM, version 3.6. available at https://gadm.org/index.html.

Population. Landscan at the 30″ by 30″ level for the year 2000, aggregated up to first-level administrative regions. Source: Bright and Coleman (2001).

Income. G-Econ 4.0 at the 1° by 1° level for the year 2000, aggregated up to first-level administrative regions. If a 1° by 1° grid cell overlaps more than one first-level administrative region, the income of that grid cell gets allocated to the different regions using their respective population shares, computed using Landscan population data at the 30″ by 30″ level. Source: Nordhaus et al. (2006).

Language composition. Language use data at the 5 km by 5 km level, aggregated up to first-level administrative regions. Source: Desmet, Gomes and Ortuño Ortín (2020).

Secessionist movements. Number, presence, and interest in all active secessionist movements around the world. Using Python, we scraped all the Wikipedia pages about active separatists movements in the world. For each movement we have information on its type (e.g., political party, militant, civil, or social organization,...), the link to any existing Wikipedia entries on it, and the name and location of the region it is associated with. This provides us with a list of 2,529 active secessionist movements. For each of these secessionist movements we then identify its geographical location at the first-level administrative region across countries. Specifically, we link each movement to all first-level administrative regions as defined in GADM v.36 for which it is trying to obtain autonomy, secession, or independence. Since the actual region associated with the movement may not always overlap perfectly with only one administrative region, we link it to all administrative regions that include or intersect the actual secessionist region. Whenever Wikipedia provides information on the administrative regions linked to the movement and these are the same as presented in GADM v.36, we directly use those. For all others, we use the location data of the proposed state, independent/autonomous region, or equivalent as provided by Wikipedia to identify the first-level administrative regions in GADM v.36 that intersect or contain it. Additionally, using Python, we scraped information on the number of visitors to all the Wikipedia entries of each movement in all languages over the 2015-2020 period. If a movement had no Wikipedia entries we assigned zero views to it. Source: List of active secessionist movements from Wikipedia at https://en.wikipedia.org/wiki/Lists_of_active_separatist_movements accessed on October 12, 2020. Wikipedia lists all movements active by continent (Africa, Asia, Europe, North America, Oceania, South America). Authors' computations.

Fragile State Index (FSI). Fragile States Index (FSI), calculated by the Fund for Peace. FSI is based on a conflict assessment framework that was developed by the Fund for Peace for assessing the vulnerability of states to collapse. Their framework was originally designed to measure this vulnerability and assess how it might affect projects in the field, and continues to be used widely by policy makers, field practitioners, and local community networks. The methodology uses both qualitative and quantitative indicators, relies on public source data, and produces quantifiable results. Twelve conflict risk indicators are used to measure the condition of a state at any given moment. The indicators provide a snapshot in time that can be measured against other snapshots in a time series to determine whether conditions are improving or worsening. Specifically, it is based on the following subcomponents: C1: Security Apparatus, C2: Factionalized Elites, C3: Group Grievance, E1: Economic Decline, E2: Uneven Economic Development, E3: Human Flight and Brain Drain, P1: State Legitimacy, P2: Public Services, P3: Human Rights and Rule of Law, S1: Demographic Pressures, S2: Refugees and IDPs, X1: External Intervention. Source: The data is accessible at https://fragilestatesindex.org/. Accessed on June 23, 2021.

Regional Authority Index (FAI). The Regional Authority Index (RAI) measures the authority in self rule and shared rule exercised by regional governments within their countries collected by Hooghe et al. (2016). It includes 96 countries (including China, India, Pakistan, Ukraine), has a temporal coverage from 1950 through 2018, and covers metropolitan and indigenous regions alongside conventional regions. Scoring is annual and the unit of analysis is the individual region. Regional authority is conceived as composed of self-rule (the authority exercised by a regional government over those who live in the region) and shared rule (the authority exercised by a region or its representatives in the country as a whole). Each domain is disaggregated in five dimensions that estimate fiscal, administrative, political, and constitutional authority. Source: Data available at https://garymarks.web.unc.edu/data/regional-authority-2/. Accessed on September 16, 2021.

Conflict data. Measures of the number of deaths from various types of conflict and terrorist attacks as provided by ACLED, UCDP/PRIO, and GTD at the first-level administrative regions across the world. Additionally, the number of years of conflict and number of conflict events for various conflict types at the first-level administrative regions across the world. We employ the types of conflict usually employed in the literature (Depetris-Chauvin and Özak, 2020; Moscona et al., 2020). Types of conflict based on ACLED data are: all, civil (where one of the actors must be the government), non-civil (where none of the actors can be the government), ethnic (where the conflict is driven by ethnic animosity, identified by all conflicts where the word "ethnic" appears in ACLED's actor names, notes, or associated actors), local (where only local actors are involved, this includes all conflicts where the ACLED interaction code is in $\{40, 41, \ldots, 48, 50, \ldots, 58, 60, \ldots, 67\}$). Using UCDP/PRIO we also construct measures for the following types of conflict: state-based (where one of the sides is the government), non-state based (where the government does not take part in conflict), territory (includes only conflicts over territorial disputes as identified by MILC), government (includes only conflicts over the type of political system, the replacement of the central government or the change of its composition, as identified by MILC). We use GTD to construct measures due to terrorist attacks. Sources: UCDP/PRIO Georeferenced Event Dataset (GED) v20.1(Pettersson et al. 2021). Data available at https://ucdp.uu.se/downloads/. Accessed on September 19, 2020. UCDP Managing Intrastate Low-intensity Conflict (MILC) dataset v.10 (Melander et al. 2009). Data available at https://ucdp.uu.se/downloads/. Accessed on September 19, 2020. Armed Conflict Location and Event Data-ACLED (Raleigh et al. 2010). Data available at https://acleddata.com/data-export-tool/. Accessed on September 17, 2020. Global Terrorism Database (GTD). GTD is available at https://www.start.umd.edu/gtd/. Analysis uses the September 2019 version accessed on November 6, 2019. Authors' computations.

Additional data references.

Bright, E. and P. R. Coleman, 2001. *Landscan 2000,* Oak Ridge, TN: Oak Ridge National Laboratory.

Depetris-Chauvin, E., and Ö. Özak, 2020. "Borderline Disorder: (De facto) Historical Ethnic Borders and Contemporary Conflict in Africa," unpublished manuscript.

Nordhaus, W., Azam, Q., Novoa Corderi, D., Hood, K., Victor, N., Mohammed, M., Miltner, A. and J. Weiss, 2006. "The G-Econ Database on Gridded Output: Methods and Data," New Haven, CT: Yale University.

Melander, E., Möller, F. and M. Öberg, 2009. "Managing Intrastate Low-Intensity Armed Conflict 1993-2004: A New Dataset," *International Interactions*, 35(1): 58-85.

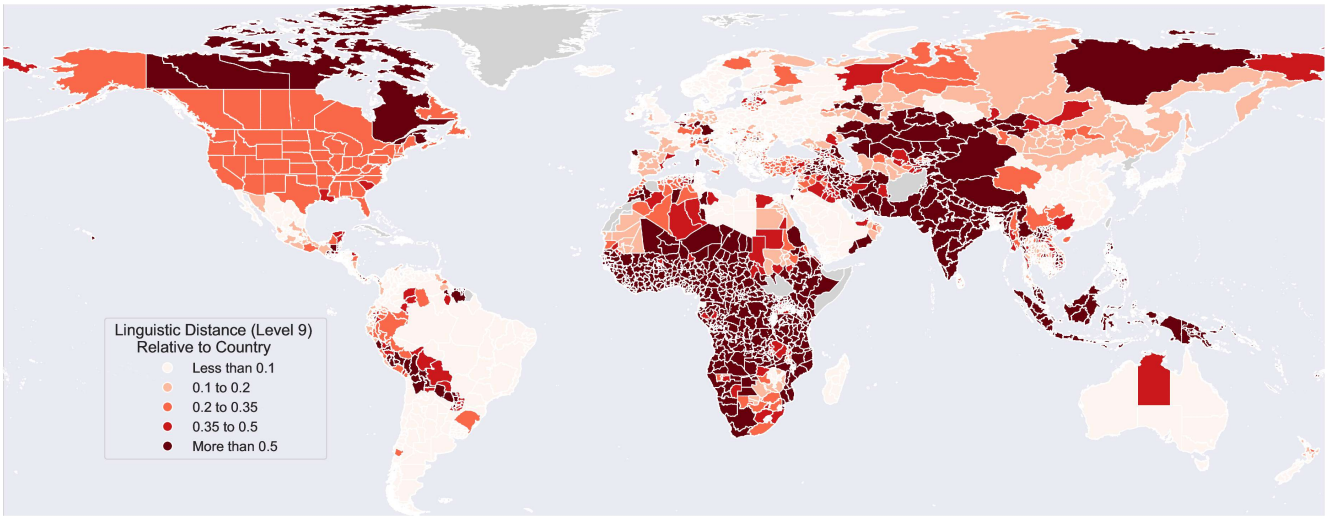
Moscona, J., Nunn, N. and J. A. Robinson, 2020. "Segmentary Lineage Organization and Conflict in Sub?Saharan Africa," *Econometrica*, 88(5): 1999-2036.

Pettersson, T., Davis, S., Deniz, A., Engström, G., Hawach, N., Högbladh, S., Sollenberg, M. and M. Öberg, 2021. "Organized Violence 1989-2020, with a Special Emphasis on Syria," *Journal of Peace Research*, 58(4): 809-825.

B Additional Tables and Figures

B.1 Linguistic Distances at Different Levels of Aggregation

(a) Aggregation Level 9



(b) Aggregation Level 2

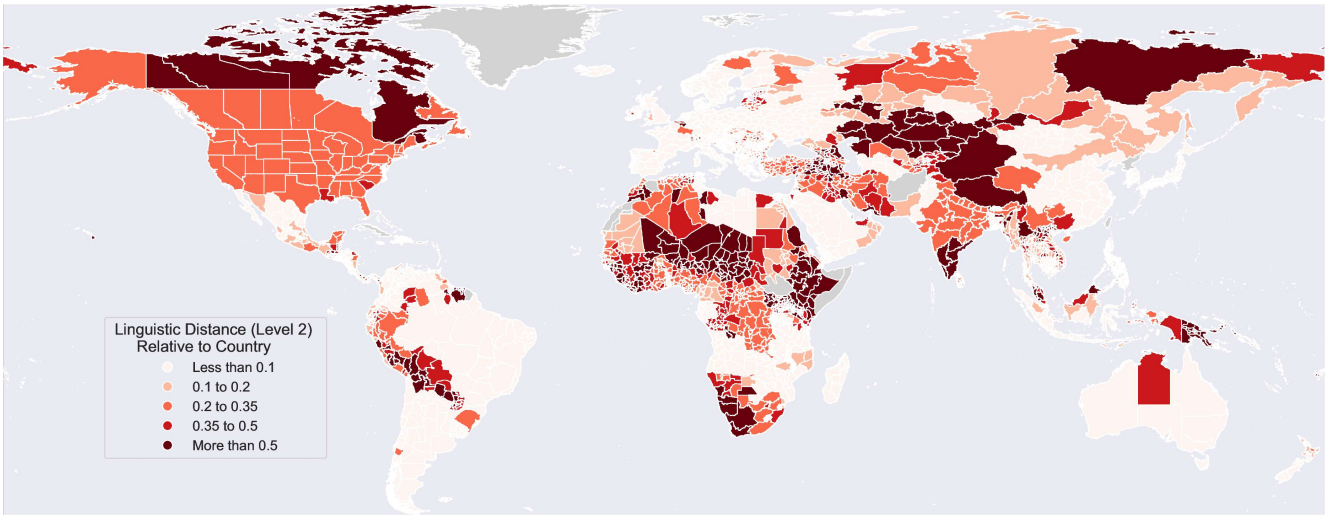
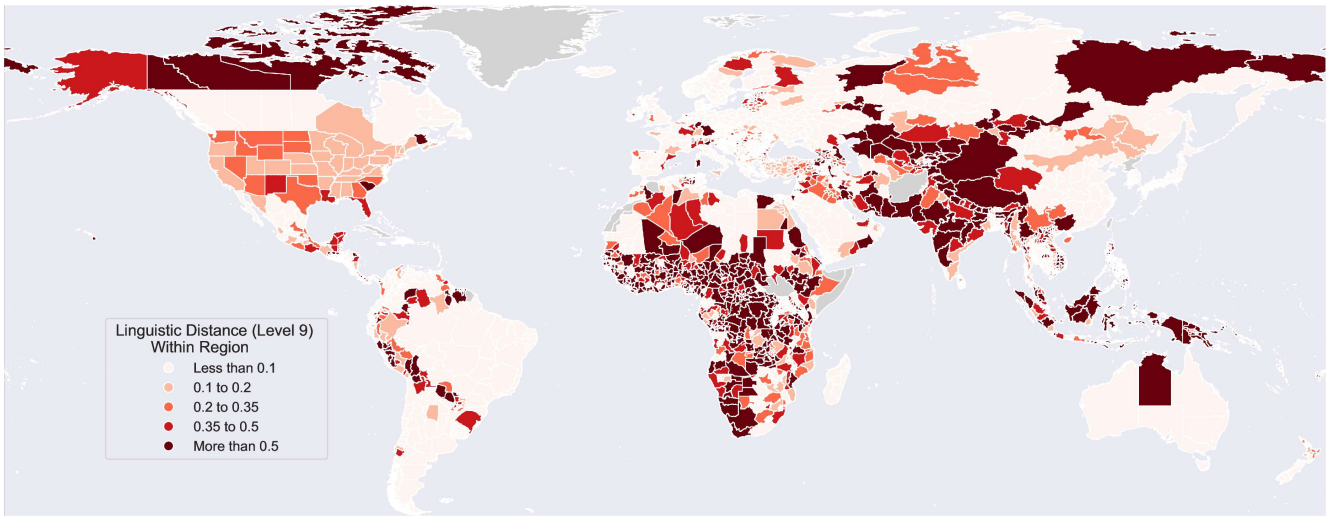


Figure B.1: Linguistic Distance between Subnational Regions and Countries: Levels of Aggregation 9 and 2

(a) Aggregation Level 9



(b) Aggregation Level 2

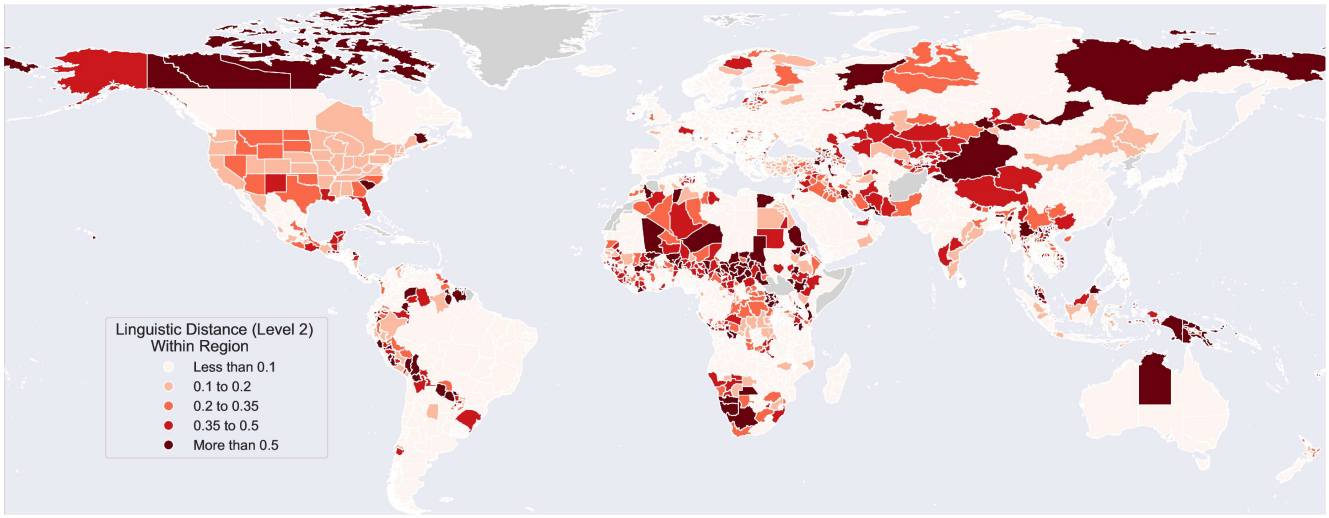


Figure B.2: Within-Subnational Region Linguistic Distance: Levels of Aggregation 9 and 2

B.2 Propensity to Secede by Country: Baseline Calibration

Table B.1: Propensity to Secede by Country: Baseline Calibration

		Population		Share			Population		Share			Population		Share
		Nat	Reg	Reg			Nat	Reg	Reg			Nat	Reg	Reg
1	Micronesia	0.678	0.702	0.500	59	Belize	0.059	0.096	0.167	117	Denmark	0.005	0.000	0.000
2	South Sudan	0.632	0.561	0.500	60	Uganda	0.055	0.056	0.069	118	Bolivia	0.004	0.000	0.000
3	Comoros	0.415	0.424	0.667	61	Serbia	0.053	0.000	0.000	119	Kuwait	0.004	0.000	0.000
4	Vanuatu	0.367	0.265	0.333	62	Panama	0.052	0.011	0.077	120	Sweden	0.004	0.000	0.000
5	Papua NG	0.363	0.348	0.409	63	Philippines	0.052	0.040	0.063	121	Tunisia	0.004	0.000	0.000
6	India	0.336	0.342	0.314	64	Yemen	0.049	0.034	0.095	122	Mauritania	0.004	0.000	0.000
7	Cameroon	0.335	0.203	0.200	65	Russia	0.047	0.042	0.036	123	Macedonia	0.004	0.000	0.000
8	Togo	0.319	0.129	0.200	66	Finland	0.045	0.000	0.000	124	USA	0.004	0.004	0.020
9	Tanzania	0.319	0.310	0.333	67	Madagascar	0.045	0.000	0.000	125	Nicaragua	0.003	0.000	0.000
10	Benin	0.274	0.280	0.333	68	Syria	0.043	0.000	0.000	126	Slovakia	0.003	0.000	0.000
11	Solomon Isl	0.254	0.355	0.500	69	Kosovo	0.043	0.000	0.000	127	Norway	0.003	0.000	0.000
12	Côte d'Ivoire	0.235	0.195	0.214	70	Senegal	0.043	0.046	0.071	128	Poland	0.002	0.000	0.000
13	Chad	0.229	0.199	0.174	71	Malawi	0.043	0.031	0.074	129	Australia	0.002	0.000	0.000
14	Sudan	0.224	0.230	0.167	72	Uzbekistan	0.042	0.000	0.000	130	Sri Lanka	0.002	0.000	0.000
15	Namibia	0.202	0.184	0.154	73	Laos	0.042	0.000	0.000	131	Slovenia	0.002	0.000	0.000
16	Pakistan	0.198	0.230	0.250	74	Libya	0.041	0.040	0.091	132	Latvia	0.002	0.000	0.000
17	Nigeria	0.195	0.130	0.162	75	France	0.037	0.000	0.000	133	Chile	0.002	0.000	0.000
18	DRC	0.188	0.065	0.038	76	Vietnam	0.035	0.011	0.016	134	El Salvador	0.002	0.000	0.000
19	Oman	0.175	0.096	0.182	77	Bhutan	0.035	0.000	0.000	135	Bulgaria	0.002	0.000	0.000
20	Indonesia	0.171	0.132	0.303	78	Peru	0.034	0.000	0.000	136	Iraq	0.002	0.000	0.000
21	CAR	0.166	0.144	0.059	79	Azerbaijan	0.031	0.000	0.000	137	Trinidad	0.001	0.000	0.000
22	Rep Congo	0.159	0.021	0.083	80	Suriname	0.027	0.000	0.000	138	Albania	0.001	0.000	0.000
23	Thailand	0.158	0.152	0.078	81	Mexico	0.027	0.000	0.000	139	Bosnia	0.001	0.000	0.000
24	Burkina Faso	0.155	0.143	0.154	82	Ecuador	0.024	0.000	0.000	140	Israel	0.001	0.000	0.000
25	Nepal	0.154	0.000	0.000	83	Guyana	0.023	0.000	0.000	141	Dom Rep	0.001	0.000	0.000
26	South Africa	0.148	0.167	0.222	84	Tajikistan	0.021	0.031	0.200	142	Ireland	0.000	0.000	0.000
27	Djibouti	0.144	0.147	0.200	85	Gambia	0.020	0.000	0.000	143	Argentina	0.000	0.000	0.000
28	China	0.143	0.158	0.161	86	Gabon	0.020	0.000	0.000	144	Cyprus	0.000	0.000	0.000
29	Zimbabwe	0.140	0.058	0.100	87	Egypt	0.019	0.008	0.111	145	UAE	0.000	0.000	0.000
30	Guinea	0.140	0.190	0.125	88	Brazil	0.017	0.000	0.000	146	Haiti	0.000	0.000	0.000
31	Eq Guinea	0.138	0.003	0.143	89	Jordan	0.016	0.000	0.000	147	Antigua	0.000	0.000	0.000
32	Mali	0.135	0.000	0.000	90	Lebanon	0.015	0.000	0.000	148	Bahrain	0.000	0.000	0.000
33	Ghana	0.133	0.000	0.000	91	Algeria	0.014	0.001	0.021	149	Belarus	0.000	0.000	0.000
34	Saudi Arabia	0.131	0.135	0.077	92	Honduras	0.014	0.001	0.056	150	Burundi	0.000	0.000	0.000
35	Bangladesh	0.126	0.179	0.286	93	Costa Rica	0.014	0.000	0.000	151	Cape Verde	0.000	0.000	0.000
36	Italy	0.126	0.027	0.050	94	Colombia	0.014	0.005	0.031	152	Estonia	0.000	0.000	0.000
37	Eritrea	0.122	0.197	0.167	95	Croatia	0.014	0.000	0.000	153	Grenada	0.000	0.000	0.000
38	Myanmar	0.120	0.095	0.267	96	Mongolia	0.013	0.000	0.000	154	Iceland	0.000	0.000	0.000
39	Guinea-Bissau	0.116	0.148	0.222	97	Cambodia	0.012	0.011	0.080	155	Jamaica	0.000	0.000	0.000
40	Iran	0.115	0.143	0.129	98	Hungary	0.011	0.000	0.000	156	Kazakhstan	0.000	0.000	0.000
41	Ethiopia	0.114	0.033	0.273	99	Romania	0.011	0.000	0.000	157	Lithuania	0.000	0.000	0.000
42	Germany	0.112	0.013	0.063	100	UK	0.010	0.000	0.000	158	Luxembourg	0.000	0.000	0.000
43	Kenya	0.109	0.132	0.128	101	Venezuela	0.010	0.000	0.000	159	Malta	0.000	0.000	0.000
44	Netherlands	0.109	0.164	0.286	102	Turkey	0.010	0.000	0.000	160	Mauritius	0.000	0.000	0.000
45	Spain	0.108	0.061	0.056	103	Czech Rep	0.010	0.000	0.000	161	Montenegro	0.000	0.000	0.000
46	Sierra Leone	0.097	0.104	0.250	104	Ukraine	0.009	0.000	0.000	162	Morocco	0.000	0.000	0.000
47	Belgium	0.085	0.000	0.000	105	Moldova	0.009	0.000	0.000	163	New Zealand	0.000	0.000	0.000
48	Mozambique	0.083	0.025	0.091	106	Kyrgyzstan	0.009	0.000	0.000	164	Rwanda	0.000	0.000	0.000
49	Georgia	0.079	0.000	0.000	107	Austria	0.009	0.000	0.000	165	St Kitts	0.000	0.000	0.000
50	Brunei	0.073	0.000	0.000	108	Greece	0.009	0.000	0.000	166	St Lucia	0.000	0.000	0.000
51	Angola	0.073	0.016	0.056	109	Paraguay	0.008	0.009	0.056	167	St Vincent	0.000	0.000	0.000
52	Fiji	0.073	0.000	0.000	110	Japan	0.007	0.008	0.021	168	Samoa	0.000	0.000	0.000
53	Malaysia	0.070	0.000	0.000	111	Turkmenistan	0.007	0.000	0.000	169	Seychelles	0.000	0.000	0.000
54	Liberia	0.069	0.066	0.133	112	Switzerland	0.007	0.000	0.000	170	South Korea	0.000	0.000	0.000
55	Zambia	0.067	0.000	0.000	113	Lesotho	0.007	0.000	0.000	171	Swaziland	0.000	0.000	0.000
56	Botswana	0.066	0.007	0.067	114	Portugal	0.006	0.000	0.000	172	Tonga	0.000	0.000	0.000
57	Guatemala	0.065	0.000	0.000	115	Armenia	0.006	0.000	0.000	173	Uruguay	0.000	0.000	0.000
58	Niger	0.059	0.029	0.125	116	Canada	0.006	0.001	0.077					

This table provides for each country three measures of the propensity to secede: the share of the country's population in favor of secession, the share of the country's population living in regions with a majority in favor of secession, and the share of the country's regions with a majority in favor of secession. These predicted measures are based on the alternative calibration. Abbreviations used: Antigua: Antigua and Barbuda; Bosnia: Bosnia and Herzegovina; CAR: Central African Republic; Czech Rep: Czech Republic; DRC: Democratic Republic of the Congo; Dom Rep: Dominican Republic; Eq Ginea: Equatorial Guinea; Papua NG: Papua New Guinea; Rep Congo: Republic of Congo; Solomon Isl: Solomon Islands; St Kitts: Saint Kitts and Nevis; St Vincent: Saint Vincent and the Grenadines; Trinidad: Trinidad and Tobago; UAE: United Arab Emirates.

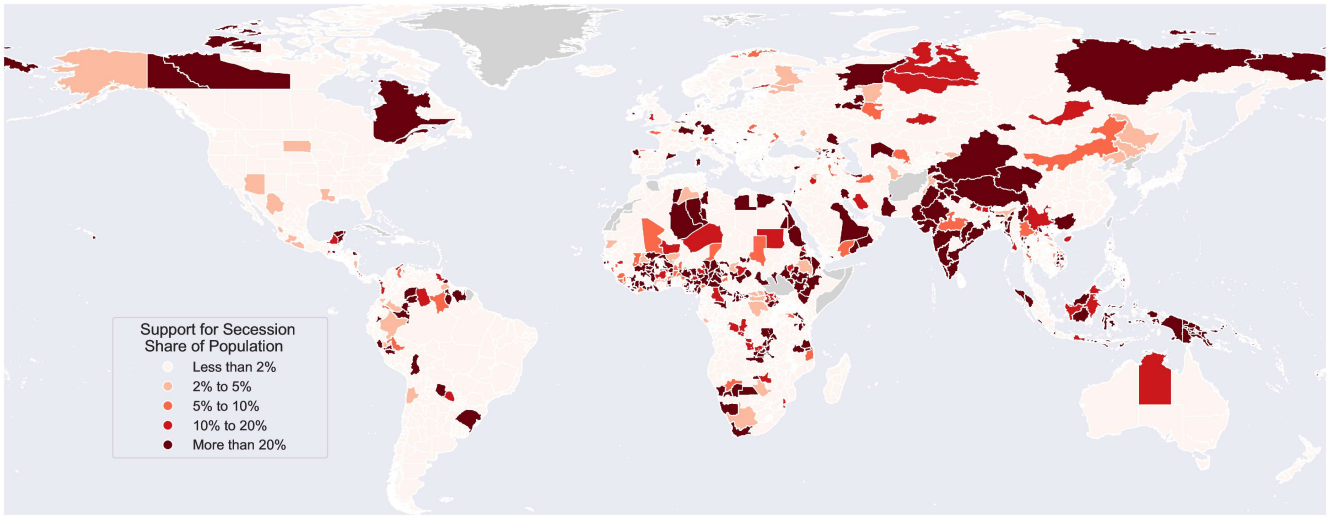
B.3 Propensity to Secede by Country: Alternative Calibration

Table B.2: Propensity to Secede by Country: Alternative Calibration

		Population		Share			Population		Share			Population		Share
		Nat	Reg	Reg			Nat	Reg	Reg			Nat	Reg	Reg
1	Pakistan	0.674	0.767	0.375	59	Kenya	0.037	0.036	0.043	117	USA	0.003	0.004	0.020
2	Papua NG	0.569	0.622	0.636	60	France	0.036	0.000	0.000	118	Macedonia	0.003	0.000	0.000
3	Belgium	0.560	0.590	0.333	61	Yemen	0.035	0.034	0.095	119	Greece	0.003	0.000	0.000
4	India	0.458	0.496	0.400	62	China	0.033	0.017	0.065	120	Norway	0.003	0.000	0.000
5	Ethiopia	0.456	0.640	0.545	63	Azerbaijan	0.032	0.000	0.000	121	Comoros	0.002	0.000	0.000
6	South Sudan	0.381	0.373	0.500	64	Mozambique	0.030	0.000	0.000	122	Nicaragua	0.002	0.000	0.000
7	Togo	0.362	0.330	0.200	65	Peru	0.029	0.000	0.000	123	Sri Lanka	0.002	0.000	0.000
8	Vanuatu	0.301	0.240	0.333	66	Suriname	0.027	0.000	0.000	124	Australia	0.002	0.000	0.000
9	Liberia	0.298	0.325	0.333	67	Zambia	0.026	0.000	0.000	125	Gabon	0.002	0.000	0.000
10	Burkina Faso	0.268	0.196	0.231	68	Guinea-Bissau	0.026	0.011	0.111	126	Latvia	0.002	0.000	0.000
11	Côte d'Ivoire	0.260	0.241	0.214	69	Vietnam	0.025	0.011	0.016	127	El Salvador	0.002	0.000	0.000
12	Chad	0.239	0.199	0.130	70	Serbia	0.025	0.000	0.000	128	Chile	0.002	0.000	0.000
13	Canada	0.233	0.236	0.154	71	Eq Guinea	0.024	0.003	0.143	129	Slovenia	0.002	0.000	0.000
14	CAR	0.217	0.305	0.235	72	Botswana	0.021	0.007	0.067	130	Trinidad	0.001	0.000	0.000
15	Philippines	0.215	0.196	0.100	73	Ireland	0.021	0.000	0.000	131	Bosnia	0.001	0.000	0.000
16	Namibia	0.210	0.184	0.154	74	Tajikistan	0.021	0.031	0.200	132	Albania	0.001	0.000	0.000
17	Nigeria	0.193	0.188	0.216	75	Italy	0.021	0.027	0.050	133	Israel	0.001	0.000	0.000
18	Indonesia	0.174	0.140	0.212	76	Guyana	0.020	0.015	0.100	134	Dom Rep	0.001	0.000	0.000
19	Mali	0.174	0.000	0.000	77	Brazil	0.017	0.000	0.000	135	Bulgaria	0.001	0.000	0.000
20	Thailand	0.158	0.152	0.078	78	Syria	0.015	0.000	0.000	136	Sweden	0.001	0.000	0.000
21	Benin	0.150	0.184	0.250	79	Lebanon	0.015	0.000	0.000	137	Slovakia	0.001	0.000	0.000
22	Djibouti	0.144	0.147	0.200	80	Honduras	0.014	0.001	0.056	138	Poland	0.000	0.000	0.000
23	Georgia	0.144	0.098	0.083	81	Rep Congo	0.013	0.021	0.083	139	Cyprus	0.000	0.000	0.000
24	Bolivia	0.142	0.000	0.000	82	Mexico	0.013	0.000	0.000	140	Argentina	0.000	0.000	0.000
25	Nepal	0.138	0.000	0.000	83	Mongolia	0.013	0.000	0.000	141	Angola	0.000	0.000	0.000
26	Sierra Leone	0.136	0.104	0.250	84	Ukraine	0.012	0.000	0.000	142	Antigua	0.000	0.000	0.000
27	Cameroon	0.136	0.000	0.000	85	Turkey	0.012	0.017	0.012	143	Bahrain	0.000	0.000	0.000
28	Saudi Arabia	0.135	0.135	0.077	86	Costa Rica	0.012	0.000	0.000	144	Belarus	0.000	0.000	0.000
29	Solomon Isl	0.131	0.192	0.300	87	Kazakhstan	0.011	0.000	0.000	145	Burundi	0.000	0.000	0.000
30	Oman	0.122	0.000	0.000	88	Niger	0.011	0.000	0.000	146	Cape Verde	0.000	0.000	0.000
31	Eritrea	0.122	0.197	0.167	89	Croatia	0.011	0.000	0.000	147	Estonia	0.000	0.000	0.000
32	Guinea	0.113	0.190	0.125	90	Egypt	0.010	0.004	0.074	148	Grenada	0.000	0.000	0.000
33	Germany	0.112	0.013	0.063	91	Romania	0.010	0.000	0.000	149	Haiti	0.000	0.000	0.000
34	Spain	0.108	0.061	0.056	92	Venezuela	0.010	0.000	0.000	150	Iceland	0.000	0.000	0.000
35	DRC	0.101	0.085	0.077	93	Cambodia	0.010	0.011	0.080	151	Iraq	0.000	0.000	0.000
36	Iran	0.094	0.101	0.065	94	Moldova	0.009	0.000	0.000	152	Jamaica	0.000	0.000	0.000
37	Netherlands	0.094	0.144	0.214	95	Kyrgyzstan	0.009	0.000	0.000	153	Jordan	0.000	0.000	0.000
38	Tanzania	0.092	0.055	0.067	96	UK	0.009	0.000	0.000	154	Lesotho	0.000	0.000	0.000
39	Guatemala	0.091	0.068	0.045	97	Algeria	0.009	0.001	0.021	155	Lithuania	0.000	0.000	0.000
40	Belize	0.083	0.096	0.167	98	Malawi	0.008	0.012	0.037	156	Luxembourg	0.000	0.000	0.000
41	Fiji	0.078	0.000	0.000	99	Colombia	0.008	0.000	0.000	157	Madagascar	0.000	0.000	0.000
42	Myanmar	0.077	0.042	0.200	100	Zimbabwe	0.008	0.000	0.000	158	Malta	0.000	0.000	0.000
43	Brunei	0.073	0.000	0.000	101	Portugal	0.008	0.000	0.000	159	Mauritius	0.000	0.000	0.000
44	Micronesia	0.072	0.000	0.000	102	Czech Rep	0.008	0.000	0.000	160	Montenegro	0.000	0.000	0.000
45	Ghana	0.069	0.000	0.000	103	Japan	0.007	0.008	0.021	161	Morocco	0.000	0.000	0.000
46	South Africa	0.068	0.100	0.111	104	Turkmenistan	0.007	0.000	0.000	162	New Zealand	0.000	0.000	0.000
47	Gambia	0.064	0.000	0.000	105	Paraguay	0.007	0.009	0.056	163	Rwanda	0.000	0.000	0.000
48	Malaysia	0.058	0.000	0.000	106	Bangladesh	0.007	0.000	0.000	164	St Kitts	0.000	0.000	0.000
49	Libya	0.056	0.058	0.136	107	Austria	0.006	0.000	0.000	165	Saint Lucia	0.000	0.000	0.000
50	Senegal	0.052	0.046	0.071	108	Armenia	0.006	0.000	0.000	166	St Vincent	0.000	0.000	0.000
51	Panama	0.049	0.011	0.077	109	Switzerland	0.005	0.000	0.000	167	Samoa	0.000	0.000	0.000
52	Russia	0.048	0.061	0.072	110	Denmark	0.005	0.000	0.000	168	Seychelles	0.000	0.000	0.000
53	Laos	0.048	0.000	0.000	111	Hungary	0.004	0.000	0.000	169	South Korea	0.000	0.000	0.000
54	Sudan	0.045	0.058	0.111	112	Kuwait	0.004	0.000	0.000	170	Swaziland	0.000	0.000	0.000
55	Kosovo	0.043	0.000	0.000	113	Mauritania	0.004	0.000	0.000	171	Tonga	0.000	0.000	0.000
56	Uzbekistan	0.043	0.058	0.071	114	Tunisia	0.004	0.000	0.000	172	UAE	0.000	0.000	0.000
57	Uganda	0.040	0.029	0.052	115	Ecuador	0.004	0.000	0.000	173	Uruguay	0.000	0.000	0.000
58	Finland	0.038	0.000	0.000	116	Bhutan	0.004	0.000	0.000					

This table provides for each country three measures of the propensity to secede: the share of the country's population in favor of secession, the share of the country's population living in regions with a majority in favor of secession, and the share of the country's regions with a majority in favor of secession. These predicted measures are based on the baseline calibration. Abbreviations used: Antigua: Antigua and Barbuda; Bosnia: Bosnia and Herzegovina; CAR: Central African Republic; Czech Rep: Czech Republic; DRC: Democratic Republic of the Congo; Dom Rep: Dominican Republic; Eq Ginea: Equatorial Guinea; Papua NG: Papua New Guinea; Rep Congo: Republic of Congo; Solomon Isl: Solomon Islands; St Kitts: Saint Kitts and Nevis; St Vincent: Saint Vincent and the Grenadines; Trinidad: Trinidad and Tobago; UAE: United Arab Emirates.

(a) Share of Regional Population in Favor of Secession: Alternative Calibration



(b) Share of Country Population in Favor of Secession: Alternative Calibration

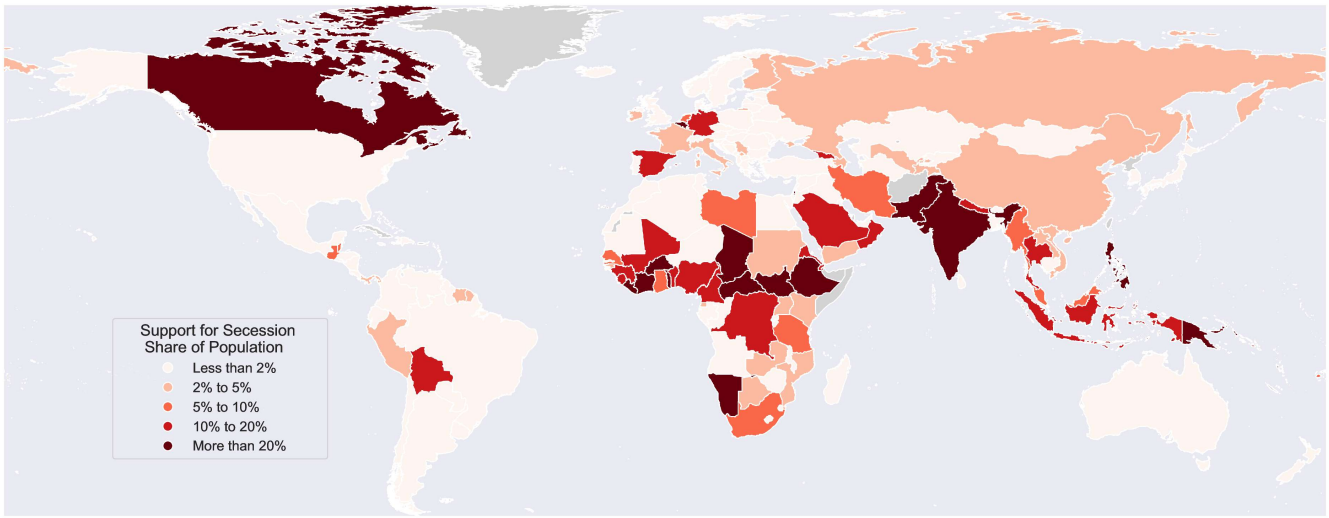


Figure B.3: Share of Population in Favor of Secession: Alternative Calibration

B.4 Accuracy of Model Predictions: Additional Results

Table B.3: Predicted Secession and Contemporary Secessionist Activity

Panel A: Sambanis et al.

	Secessionist Activity in Country			Log[1+# Secessionist Groups in Country]		
	(1)	(2)	(3)	(4)	(5)	(6)
Log[1 + # Secessionist Regions]	0.197***			0.961***		
	(0.059)			(0.208)		
Log[1 + Share Pop. Pro-Secession]		1.670***			6.334***	
		(0.526)			(1.869)	
At least 1 Region Pro-Secession			0.267***			1.187***
			(0.081)			(0.268)
Log[1+# Regions]	Yes	Yes	Yes	Yes	Yes	Yes
WB Region FE	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted-R^2	0.13	0.16	0.14	0.24	0.23	0.23
Observations	173	173	173	173	173	173

Panel B: Sambanis et al. & Wikipedia

	Secessionist Activity in Country			Log[1+# Secessionist Groups in Country]		
	(1)	(2)	(3)	(4)	(5)	(6)
Log[1 + # Secessionist Regions]	0.201***			0.943***		
	(0.059)			(0.214)		
Log[1 + Share Pop. Pro-Secession]		1.682***			6.591***	
		(0.556)			(1.931)	
At least 1 Region Pro-Secession			0.303***			1.266***
			(0.080)			(0.268)
Log[1+# Regions]	Yes	Yes	Yes	Yes	Yes	Yes
WB Region FE	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted-R^2	0.12	0.14	0.14	0.25	0.26	0.26
Observations	173	173	173	173	173	173

Notes: Heteroskedasticity robust standard error estimates clustered at the country-level are reported in parentheses; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests.

Table B.4: Predicted Demand for Secession and Fragile State Index Subcomponents

	Fragile State Index Subcomponents (2006-2021)											
	C1	C2	C3	E1	E2	E3	P1	P2	P3	S1	S2	X1
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Log[1 + Share Pop. Pro-Secession]	5.940***	4.164**	4.173*	3.354***	4.630***	2.587	3.612**	5.627***	2.174	3.729***	3.840*	2.918
	(1.908)	(1.892)	(2.319)	(1.253)	(1.279)	(1.703)	(1.749)	(1.504)	(2.206)	(1.256)	(2.128)	(1.824)
WB Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean Dependent	5.71	6.45	6.24	5.65	5.88	5.67	6.05	5.50	5.75	5.76	5.32	5.75
Adjusted-R^2	0.34	0.21	0.17	0.40	0.50	0.39	0.27	0.59	0.34	0.62	0.34	0.31
Observations	167	167	167	167	167	167	167	167	167	167	167	167

Notes: Heteroskedasticity robust standard error estimates clustered at the country-level are reported in parentheses; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests. Each column shows one subcomponent of the Fragile State Index. C1: Security Apparatus, C2: Factionalized Elites, C3: Group Grievance, E1: Economic Decline, E2: Uneven Economic Development, E3: Human Flight and Brain Drain, P1: State Legitimacy, P2: Public Services, P3: Human Rights and Rule of Law, S1: Demographic Pressures, S2: Refugees and IDPs, X1: External Intervention.

Table B.5: Predicted Demand for Secession and Regional Authority Index Subcomponents

	Regional Authority Index (1950-2016)											
	Self-Rule						Shared-Rule					
	SE1	SE2	SE3	SE4	SE5	SE	SH1	SH2	SH3	SH4	SH5	SH
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Log[1 + Share Pop.	10.715***	7.500***	5.149*	2.514	14.025***	39.820***	2.621*	1.342	1.376	0.091	8.292**	13.706*
Pro-Secession]	(3.729)	(2.674)	(2.819)	(2.141)	(4.445)	(14.625)	(1.502)	(1.251)	(1.227)	(0.754)	(3.958)	(8.054)
WB Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean Dependent	2.21	1.70	1.06	1.02	3.04	9.03	0.38	0.31	0.28	0.18	0.85	2.00
Adjusted-R^2	0.31	0.21	0.22	0.19	0.28	0.25	0.02	0.19	0.01	-0.02	0.17	0.09
Observations	89	89	89	89	89	89	89	89	89	89	89	89

Notes: Heteroskedasticity robust standard error estimates clustered at the country-level are reported in parentheses; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests. Each column shows one subcomponent of the Regiuonal Autonomy Index. First 6 columns contain subcomponents measuring the level of self-rule within regions, the last 6 columns measure the level of shared power between regions and central governments. SE1: The extent to which a regional government is autonomous rather than deconcentrated. SE2: The range of policies for which a regional government is responsible. SE3: The extent to which a regional government can independently tax its population. SE4: The extent to which a regional government can borrow. SE5: The extent to which a region has an independent legislature and executive, which is the sum of assembly and executive. SE: The authority exercised by a regional government over those who live in the region, which is the sum of SE1-SE5. SH1:The extent to which regional representatives co–determine national legislation. SH2: The extent to which a regional government co–determines national policy in intergovernmental meetings. SH3: The extent to which regional representatives co–determine the distribution of national tax revenues. SH4: The extent to which a regional government co–determines subnational and national borrowing constraints. SH5: The extent to which regional representatives co–determine constitutional change. SH: The authority exercised by a regional government or its representatives in the country as a whole, which is teh sum of SH1-SH5.

B.5 Accuracy of Model Predictions: Alternative Calibration

Table B.6: Predicted Demand for Secession and Contemporary Secessionist Activity: Alternative Calibration

Panel A: Country-Level Analysis	Interest in Secession								
	Secessionist Activity in Country						Log[Number of Wikipedia Page Views]		
	Log[1+# Secessionist Regions]			Log[1+# Secessionist Groups]			Secessionist Organization		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Log[1 + # Secessionist Regions]	0.336***			0.562***			2.293**		
	(0.117)			(0.212)			(0.943)		
Log[1 + Share Pop. Pro-Secession]		1.519**			2.724**			13.645**	
		(0.627)			(1.224)			(6.063)	
At least 1 Region Pro-Secession			0.417***			0.700***			3.038***
			(0.136)			(0.241)			(1.093)
WB Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Log[1+# Regions]	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted-R^2	0.33	0.31	0.33	0.27	0.25	0.27	0.19	0.18	0.19
Observations	173	173	173	173	173	173	173	173	173

Panel B: Regional-Level Analysis	Interest in Secession								
	Secessionist Activity in Region						Log[Total Number of Wikipedia Page Views]		
	Secessionist Region			Log[1+# Secessionist Groups in Region]			Secessionist Organization		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Log[1 + Sh. Pop. Pro-Secession]	0.275***			0.431***			2.870***		
	(0.066)			(0.109)			(0.712)		
At least 10% Pro-Secession		0.140***			0.210***			1.332***	
		(0.026)			(0.041)			(0.244)	
At least 50% Pro-Secession			0.117***			0.192***			1.323***
			(0.037)			(0.065)			(0.472)
Country FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted-R^2	0.21	0.22	0.21	0.24	0.24	0.24	0.22	0.22	0.21
Observations	3003	3003	3003	3003	3003	3003	3003	3003	3003

Notes: Heteroskedasticity robust standard error estimates clustered at the country-level are reported in parentheses; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests.

Table B.7: Predicted Demand for Secession and Contemporary Secessionist Activity: Alternative Calibration

Panel A: Sambanis et al.

	Secessionist Activity in Country			Log[1+# Secessionist Groups in Country]		
	(1)	(2)	(3)	(4)	(5)	(6)
Log[1 + # Secessionist Regions]	0.119**			0.738***		
	(0.060)			(0.220)		
Log[1 + Share Pop. Pro-Secession]		0.950***			4.413***	
		(0.352)			(1.316)	
At Least 1 Region Pro-Secession			0.173**			0.864***
			(0.078)			(0.269)
WB Region FE	Yes	Yes	Yes	Yes	Yes	Yes
Log[1+# Regions]	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted-R^2	0.10	0.12	0.11	0.21	0.20	0.20
Observations	173	173	173	173	173	173

Panel B: Sambanis et al. & Wikipedia

	Secessionist Activity in Country			Log[1+# Secessionist Groups in Country]		
	(1)	(2)	(3)	(4)	(5)	(6)
Log[1 + # Secessionist Regions]	0.147**			0.812***		
	(0.059)			(0.221)		
Log[1 + Share Pop. Pro-Secession]		0.967***			4.684***	
		(0.359)			(1.337)	
At Least 1 Region Pro-Secession			0.219***			1.020***
			(0.076)			(0.264)
WB Region FE	Yes	Yes	Yes	Yes	Yes	Yes
Log[1+# Regions]	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted-R^2	0.09	0.10	0.11	0.23	0.22	0.24
Observations	173	173	173	173	173	173

Notes: Heteroskedasticity robust standard error estimates clustered at the country-level are reported in parentheses; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests.

Table B.8: Predicted Secession and State Fragility, Regional Autonomy, and Conflict: Alternative Calibration

Panel A: Institutional	Fragile State Index (2006-2021)						Regional Authority Index (1950-2016)		
	Rank			Index			Total		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Log[1 + # Secessionist Regions]	20.999***			9.011***			4.773**		
	(4.876)			(2.163)			(1.831)		
Log[1 + Share of Population Pro-Secession]		80.210**			32.287**			38.203***	
		(33.495)			(15.959)			(12.572)	
At least 1 Region Supports Secession			24.139***			10.263***			6.766**
			(6.483)			(2.936)			(2.864)
WB Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted-R^2	0.47	0.42	0.45	0.45	0.41	0.44	0.15	0.20	0.16
Observations	167	167	167	167	167	167	89	89	89

Panel B: Conflict	Intensity of Conflict (1997-2020)								
	Log[1 + # Deaths]			Log[1 + # Events]			Log[1 + # Years]		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Log[1 + # Secessionist Regions]	1.970***			1.025***			0.142**		
	(0.337)			(0.195)			(0.056)		
Log[1 + Share of Population Pro-Secession]		8.989***			4.380***			0.339	
		(2.546)			(1.455)			(0.266)	
At least 1 Region Supports Secession			1.864***			0.896***			0.195**
			(0.518)			(0.332)			(0.076)
WB Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted-R^2	0.43	0.36	0.38	0.19	0.14	0.15	0.88	0.87	0.88
Observations	131	131	131	131	131	131	131	131	131

Notes: Heteroskedasticity robust standard error estimates clustered at the country-level are reported in parentheses; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests.

Table B.9: Predicted Secession and Fragile State Index Subcomponents: Alternative Calibration

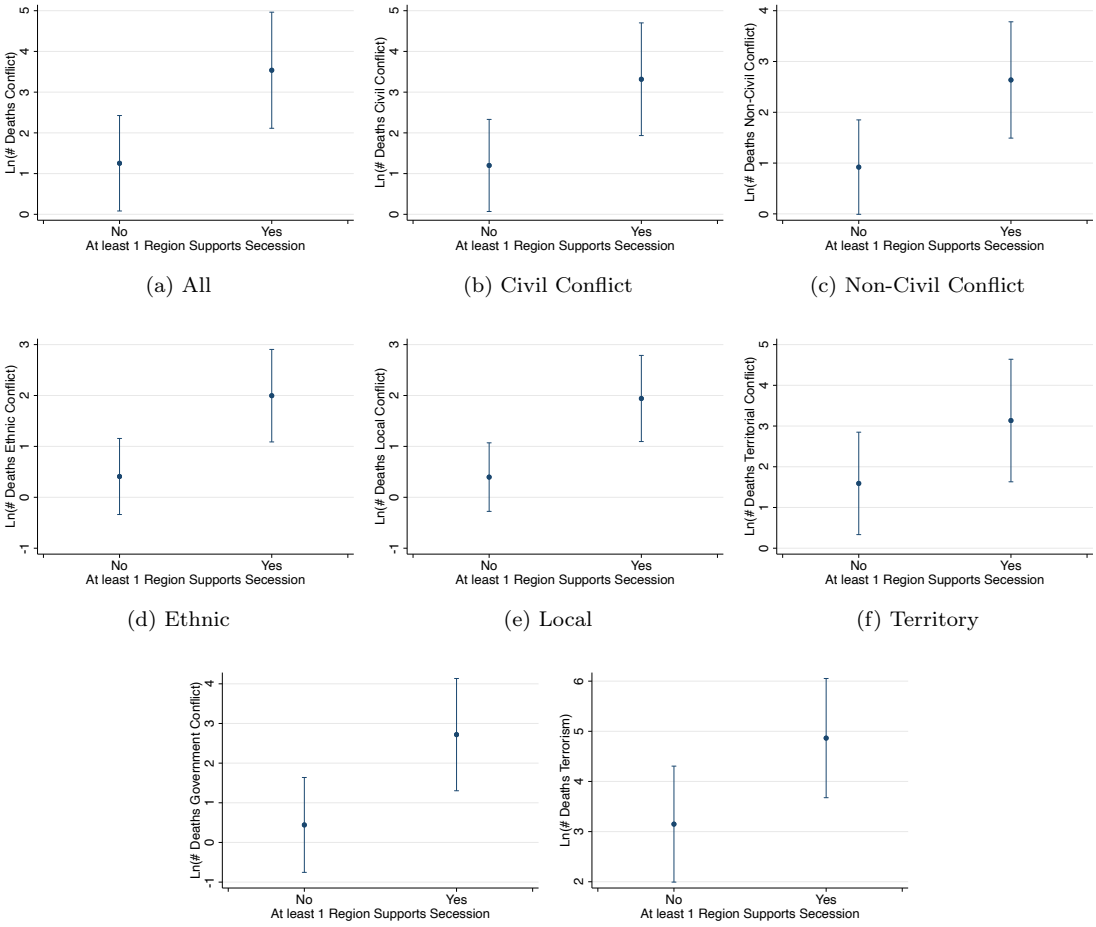
	Fragile State Index Subcomponents (2006-2021)											
	C1	C2	C3	E1	E2	E3	P1	P2	P3	S1	S2	X1
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Log[1 + # Secessionist Regions]	1.193***	0.909***	1.069***	0.373**	0.726***	0.278	0.676**	0.852***	0.809***	0.776***	0.989***	0.357
	(0.229)	(0.254)	(0.225)	(0.178)	(0.163)	(0.184)	(0.262)	(0.195)	(0.255)	(0.159)	(0.257)	(0.228)
WB Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted-R^2	0.40	0.24	0.25	0.40	0.52	0.39	0.28	0.61	0.38	0.65	0.39	0.31
Observations	167	167	167	167	167	167	167	167	167	167	167	167

Notes: Heteroskedasticity robust standard error estimates clustered at the country-level are reported in parentheses; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests. Each column shows one subcomponent of the Fragile State Index. C1: Security Apparatus, C2: Factionalized Elites, C3: Group Grievance, E1: Economic Decline, E2: Uneven Economic Development, E3: Human Flight and Brain Drain, P1: State Legitimacy, P2: Public Services, P3: Human Rights and Rule of Law, S1: Demographic Pressures, S2: Refugees and IDPs, X1: External Intervention.

Table B.10: Predicted Secession and Regional Autonomy Index Subcomponents: Alternative Calibration

	Regional Autonomy Index (1950-2016)											
	Self-Rule						Shared-Rule					
	SE1	SE2	SE3	SE4	SE5	SE	SH1	SH2	SH3	SH4	SH5	SH
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Log[1 + # Secessionist Regions]	0.660***	0.473**	0.423*	0.154	0.934***	2.635***	0.199	0.060	0.110	-0.049	0.616*	0.934
	(0.206)	(0.206)	(0.240)	(0.144)	(0.265)	(0.940)	(0.126)	(0.095)	(0.102)	(0.059)	(0.315)	(0.633)
WB Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted-R^2	0.22	0.17	0.22	0.19	0.22	0.21	0.01	0.17	0.01	-0.01	0.15	0.07
Observations	89	89	89	89	89	89	89	89	89	89	89	89

Notes: Heteroskedasticity robust standard error estimates clustered at the country-level are reported in parentheses; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests. Each column shows one subcomponent of the Regiuonal Autonomy Index. First 6 columns contain subcomponents measuring the level of self-rule within regions, the last 6 columns measure the level of shared power between regions and central governments. SE1: The extent to which a regional government is autonomous rather than deconcentrated. SE2: The range of policies for which a regional government is responsible. SE3: The extent to which a regional government can independently tax its population. SE4: The extent to which a regional government can borrow. SE5: The extent to which a region has an independent legislature and executive, which is the sum of assembly and executive. SE: The authority exercised by a regional government over those who live in the region, which is the sum of SE-SE5. SH1:The extent to which regional representatives co–determine national legislation. SH2: The extent to which a regional government co–determines national policy in intergovernmental meetings. SH3: The extent to which regional representatives co–determine the distribution of national tax revenues. SH4: The extent to which a regional government co–determines subnational and national borrowing constraints. SH5: The extent to which regional representatives co–determine constitutional change. SH: The authority exercised by a regional government or its representatives in the country as a whole, which is teh sum of SH1-SH5.



(a) All (b) Civil Conflict (c) Non-Civil Conflict

(d) Ethnic (e) Local (f) Territory

(g) Government (h) Terrorism

Figure B.4: Predicted Secession and Conflict Intensity (Deaths) across Countries: Alternative Calibration

B.6 Identity vs Income per Capita

Table B.11: Instability of Countries: Identity vs Income per Capita

			Δ Pro-Secession					Δ Pro-Secession					Δ Pro-Secession	
		Base-line	Iden Only	Inc Only			Base-line	Iden Only	Inc Only			Base-Line	Iden Only	Inc Only
1	Micronesia	0.678	0.000	-0.678	59	Belize	0.059	0.000	-0.059	117	Denmark	0.005	0.000	-0.005
2	South Sudan	0.632	-0.008	-0.632	60	Uganda	0.055	0.000	-0.055	118	Bolivia	0.004	0.137	-0.004
3	Comoros	0.415	0.000	-0.415	61	Serbia	0.053	0.000	-0.053	119	Kuwait	0.004	-0.004	-0.004
4	Vanuatu	0.367	-0.067	-0.284	62	Panama	0.052	0.000	-0.052	120	Sweden	0.004	0.000	-0.004
5	Papua NG	0.363	0.085	-0.363	63	Philippines	0.052	0.091	-0.052	121	Tunisia	0.004	0.000	-0.004
6	India	0.336	0.136	-0.336	64	Yemen	0.049	-0.029	-0.019	122	Mauritania	0.004	0.000	-0.004
7	Cameroon	0.335	-0.066	-0.335	65	Russia	0.047	0.011	-0.037	123	Macedonia	0.004	0.000	-0.004
8	Togo	0.319	0.000	-0.319	66	Finland	0.045	-0.001	-0.045	124	USA	0.004	0.000	-0.004
9	Tanzania	0.319	-0.005	-0.319	67	Madagascar	0.045	0.000	-0.045	125	Nicaragua	0.003	0.000	-0.003
10	Benin	0.274	0.000	-0.274	68	Syria	0.043	0.007	-0.043	126	Slovakia	0.003	0.000	-0.003
11	Solomon Isl	0.254	0.000	-0.254	69	Kosovo	0.043	0.000	-0.043	127	Norway	0.003	0.001	-0.003
12	Côte d'Ivoire	0.235	0.000	-0.235	70	Senegal	0.043	0.000	-0.043	128	Poland	0.002	0.002	-0.002
13	Chad	0.229	-0.034	-0.161	71	Malawi	0.043	0.000	-0.043	129	Australia	0.002	0.000	-0.002
14	Sudan	0.224	-0.180	-0.052	72	Uzbekistan	0.042	0.000	-0.042	130	Sri Lanka	0.002	0.000	-0.002
15	Namibia	0.202	-0.127	-0.057	73	Laos	0.042	0.041	-0.042	131	Slovenia	0.002	0.000	-0.002
16	Pakistan	0.198	0.033	-0.198	74	Libya	0.041	-0.036	-0.005	132	Latvia	0.002	0.073	-0.002
17	Nigeria	0.195	0.015	-0.195	75	France	0.037	-0.005	-0.037	133	Chile	0.002	0.014	-0.002
18	DRC	0.188	0.090	-0.188	76	Vietnam	0.035	0.013	-0.035	134	El Salvador	0.002	0.000	-0.002
19	Oman	0.175	0.000	-0.175	77	Bhutan	0.035	0.000	-0.035	135	Bulgaria	0.002	0.000	-0.002
20	Indonesia	0.171	0.056	-0.171	78	Peru	0.034	0.091	-0.034	136	Iraq	0.002	0.000	-0.002
21	CAR	0.166	0.135	-0.166	79	Azerbaijan	0.031	0.000	-0.031	137	Trinidad	0.001	0.000	-0.001
22	Rep Congo	0.159	0.000	-0.159	80	Suriname	0.027	0.000	-0.027	138	Albania	0.001	0.000	-0.001
23	Thailand	0.158	-0.138	0.004	81	Mexico	0.027	0.015	-0.027	139	Bosnia	0.001	0.000	-0.001
24	Burkina Faso	0.155	0.000	-0.155	82	Ecuador	0.024	0.000	-0.024	140	Israel	0.001	0.000	-0.001
25	Nepal	0.154	0.000	-0.154	83	Guyana	0.023	0.000	-0.023	141	Dom Rep	0.001	0.000	-0.001
26	South Africa	0.148	0.093	-0.148	84	Tajikistan	0.021	0.000	-0.021	142	Ireland	0.000	0.021	0.000
27	Djibouti	0.144	0.000	-0.144	85	Gambia	0.020	0.000	-0.020	143	Argentina	0.000	0.005	0.000
28	China	0.143	-0.009	-0.143	86	Gabon	0.020	0.000	-0.020	144	Cyprus	0.000	0.000	0.000
29	Zimbabwe	0.140	0.000	-0.140	87	Egypt	0.019	-0.005	-0.015	145	UAE	0.000	0.000	0.000
30	Guinea	0.140	0.000	-0.140	88	Brazil	0.017	0.000	-0.017	146	Haiti	0.000	0.000	0.000
31	Eq Guinea	0.138	0.000	-0.138	89	Jordan	0.016	0.000	-0.016	147	Antigua	0.000	0.000	0.000
32	Mali	0.135	0.000	-0.135	90	Lebanon	0.015	0.000	-0.015	148	Bahrain	0.000	0.000	0.000
33	Ghana	0.133	0.076	-0.133	91	Algeria	0.014	-0.003	-0.013	149	Belarus	0.000	0.000	0.000
34	Saudi Arabia	0.131	-0.131	0.004	92	Honduras	0.014	0.000	-0.014	150	Burundi	0.000	0.000	0.000
35	Bangladesh	0.126	-0.001	-0.126	93	Costa Rica	0.014	0.000	-0.014	151	Cape Verde	0.000	0.000	0.000
36	Italy	0.126	0.058	-0.126	94	Colombia	0.014	-0.004	-0.009	152	Estonia	0.000	0.000	0.000
37	Eritrea	0.122	0.000	-0.122	95	Croatia	0.014	-0.003	-0.014	153	Grenada	0.000	0.000	0.000
38	Myanmar	0.120	0.008	-0.120	96	Mongolia	0.013	0.001	-0.013	154	Iceland	0.000	0.000	0.000
39	Guinea-Bissau	0.116	0.000	-0.116	97	Cambodia	0.012	0.000	-0.012	155	Jamaica	0.000	0.000	0.000
40	Iran	0.115	0.063	-0.115	98	Hungary	0.011	-0.008	-0.011	156	Kazakhstan	0.000	0.017	0.000
41	Ethiopia	0.114	0.052	-0.114	99	Romania	0.011	-0.005	-0.011	157	Lithuania	0.000	0.000	0.000
42	Germany	0.112	-0.035	-0.112	100	UK	0.010	0.000	-0.010	158	Luxembourg	0.000	0.000	0.000
43	Kenya	0.109	-0.002	-0.068	101	Venezuela	0.010	0.001	-0.010	159	Malta	0.000	0.000	0.000
44	Netherlands	0.109	0.000	-0.109	102	Turkey	0.010	0.014	-0.010	160	Mauritius	0.000	0.000	0.000
45	Spain	0.108	-0.047	-0.108	103	Czech Rep	0.010	-0.001	-0.010	161	Montenegro	0.000	0.000	0.000
46	Sierra Leone	0.097	0.000	-0.097	104	Ukraine	0.009	0.016	-0.009	162	Morocco	0.000	0.001	0.000
47	Belgium	0.085	0.048	-0.085	105	Moldova	0.009	0.000	-0.009	163	New Zealand	0.000	0.000	0.000
48	Mozambique	0.083	0.111	-0.083	106	Kyrgyzstan	0.009	-0.009	-0.009	164	Rwanda	0.000	0.000	0.000
49	Georgia	0.079	0.000	-0.079	107	Austria	0.009	0.000	-0.009	165	St Kitts	0.000	0.000	0.000
50	Brunei	0.073	0.000	-0.073	108	Greece	0.009	0.041	-0.009	166	St Lucia	0.000	0.000	0.000
51	Angola	0.073	-0.002	-0.073	109	Paraguay	0.008	0.000	-0.008	167	St Vincent	0.000	0.000	0.000
52	Fiji	0.073	0.000	-0.073	110	Japan	0.007	0.000	-0.007	168	Samoa	0.000	0.000	0.000
53	Malaysia	0.070	0.001	-0.070	111	Turkmenistan	0.007	0.000	-0.007	169	Seychelles	0.000	0.000	0.000
54	Liberia	0.069	0.000	-0.069	112	Switzerland	0.007	0.000	-0.007	170	South Korea	0.000	0.000	0.000
55	Zambia	0.067	0.184	-0.067	113	Lesotho	0.007	0.000	-0.007	171	Swaziland	0.000	0.000	0.000
56	Botswana	0.066	0.041	-0.066	114	Portugal	0.006	-0.002	-0.006	172	Tonga	0.000	0.000	0.000
57	Guatemala	0.065	0.000	-0.065	115	Armenia	0.006	0.000	-0.006	173	Uruguay	0.000	0.000	0.000
58	Niger	0.059	-0.032	-0.059	116	Canada	0.006	0.229	-0.006					

This table reports for each country how support for secession depends on linguistic identity and on income per capita. Column 'Base' reports the share of the population that favors secession in the baseline calibration; column 'Iden' reports the change in that share if differences in income per capita are eliminated and only identity matters; column 'Inc' reports the change in that share if differences in identity are eliminated and only differences in income per capita matter. Antigua: Antigua and Barbuda; Bosnia: Bosnia and Herzegovina; CAR: Central African Republic; DRC: Democratic Republic of the Congo; Dom Rep: Dominican Republic; Papua NG: Papua New Guinea; Rep Congo: Republic of Congo; Solomon Isl: Solomon Islands; St Kitts: Saint Kitts and Nevis; St Vincent: Saint Vincent and the Grenadines; Trinidad: Trinidad and Tobago; UAE: United Arab Emirates.