Негосударственное образовательное учреждение
высшего профессионального образования
«Российская экономическая школа» (институт)

МАГИСТЕРСКАЯ ДИССЕРТАЦИЯ

по направлению подготовки 080100.68
Программа «Магистр экономики»

MASTER THESIS

Тема: Распад государств: случай большого федеративного государства

Title: On The Breakup of Nations: The Case of a Large Federal State

Студент/ Student:

Хазанов Алексей Аркадьевич
(Ф.И.О. полностью студента, выполнившего работу)
Alexey Khazanov
(Name of the student)

Научный руководитель/ Advisor:

Доктор Шломо Вебер, РЭШ; Доктор Пол Дауэр, РЭШ; Доктор Гюнеш Гёкмен, РЭШ; Доктор Андрей Маркевич, РЭШ.
(ученая степень, звание, место работы, Ф.И.О.)
Paul Dower, Ph.D., NES, Gunes Gokmen, Ph.D., NES, Andrei Markevich, Ph.D., NES, Shlomo Weber, Ph.D., NES.
(Name, scientific degree, scientific rank, place of work,)

Москва 2014

On The Breakup of Nations: The Case of a Large Federal State

Alexey Khazanov

Abstract

This paper presents a model of state breakup within the context of a large federal state. We consider two versions of the model: with homogeneous regions and with heterogeneity within regions. Findings suggest that geographic, income and cultural heterogeneity drive the breakup process. In this paper we suggest a possible solution for the multidimensional heterogeneity problem and we present simulations to show the existence of the equilibrium in models with quasi-linear utility functions and specific income, culture and location distributions. We also estimate the theoretical model that can be used for the empirical analysis of the large federal states' stability and we demonstrate how it can be applied.

The breakup of nations that happened in 1990-s in Yugoslavia and Soviet Union as well as the unification of Germany drew attention to possible determinants of the nation formation and breakup in economic literature.

The literature has a relatively strong consensus about the key determinants of the breakup processes, that can be divided into monetary and non-monetary factors (Desmet et al. (2011)). These are the degrees of income and cultural (ethnic, linguistic) heterogeneity respectively that lead to the differences of preferences on public good's provision. It is easy to believe that these factors are the major ones in the secession of Yugoslav republics or the clash of the Soviet Union. These factors might also play a role in separatist movements in Scotland, Basque country, Catalonia and Flanders in which a relatively high cultural dissimilarity with the rest of the corresponding countries is observed.

However, the cultural distances do not explain the reasons for the existence of influential separatist movements in regions that are of the dominant country's culture but that are geographically far from the center of the country. This effect can be observed in Greenland (Inuit Ataqatigiit party), in Alaska (Alaska Independence Party), in Russian Far East and in many other regions of the world that are isolated from the rest of the country either geographically, informationally or psychologically.

The idea of the "alienation" of some parts of the society was already examined in the literature. The conflict between "peripheral" minority groups and the country dominant ethnicity is the subject of Desmet et al. (2005) paper in which special indexes of peripheral diversity are constructed. The idea of the conflict between the "periphery" and the "center", but in another context, became the motivation for our model.

In the literature it is usually assumed (Alesina and Spolaore (1997)) that geography is highly correlated with agents' preferences which allows to simplify the models by assuming that the non-monetary heterogeneity is one-dimensional. However, Desmet et al. (2011) have shown that the partial correlation between cultural and geographic distances controlling for genetic and linguistic ones is not statistically significant. Russian Far East, where most part of the population are ethnic Russians and where the income distribution is not very different from other Russian regions[1] and where a separatist movement exists, can be a good evidence supporting their results.

This evidence became a motivation for our extension of the existing models of the nations' breakup. Our hypothesis is that the described phenomenon can be related to the **quality of information transmission** between regions and the center. This problem is especially crucial in large countries where the central government's officials can not check the situation in details and count on the information provided by local authorities. The central government (in general sense) receives a noisy signal about actual needs of the region and thus can not efficiently understand how exactly and in what amount the public goods should be provided in this region. Interview with the locals in Russian Far East[2] shows that the citizens suppose that the governance of the region from the "center" is not effective and it does not take into account the interests of the population.

We use the geographic distance proxy[3] to measure this information transmission inefficiency. The intuition behind this is that it is usually easier for a central government official to monitor the situation around the capital than to make a long trip to distant region, during which the local authorities would probably show her Potemkin villages[4][5]. From this follows that in large federal state the factor of geographic distance as a proxy for this information transmission inefficiency can matter. We construct a model for a federal state for two reasons. First, the most part of large countries (Russia, the U.S.A., Canada, Australia, India, Brazil) are federations[6]. Second,

[1] According to Federal State Statistics Service data.

[2] http://www.kommersant.ru/doc-rss/1930145, in Russian.

[3] The measure similar to Ashraf et al. (2010).

[4] In Russian history, fake villages built by Gregory Potemkin, Russian noble and the favourite of Catherine II, to show to the Russian Empress that he was successful in rebuilding the region devastated after war.

[5] A simple model that describes the game of the local authorities and the central government is presented in the Appendix C.

[6] There can be mixed types as it is in Russia, where part of the regions are republics with high degree of independence, but most part of regions are "Oblast'" that do not have this.

we want the regions to potentially have right to secede without taking into account the interests of the rest of the country. That is necessary for the equilibrium concept that is used in our model and it simplifies the calculations a lot[7].

In our paper we present a model of a country in which the borders of the regions are fixed and migration is not allowed, the regions are considered as the members of the coalition (country) and the further path of the regions after secession is not examined. Given this, our approach is closely related to Bolton and Roland (1997) and Desmet et al. (2011).

The main contribution of our paper is the construction of a tool applicable to the analysis of the stability of countries which takes into account the multidimensional heterogeneity of the regions' population. We suggest one possible solution for the problem of within-region multidimensional heterogeneity which complicates the empirical analysis of the break ups. The solution is the multilevel system of voting on optimal tax rate[8] that we call *community voting*[9] which unambiguously determines the characteristics of the decision making agents in the region. The idea to divide the regions' population into groups according to their culture, language or religion and to assume that each of these artificial communities has voters in a regional parliament. This approach is a simplification necessary for the empirical studies that consider multidimensionally heterogeneous regions and countries.

The paper is organized as follows. The first part is devoted to the literature review. In the second part we present a basic model, discuss assumptions that we make, introduce the equilibrium concept. In the third part we present the extension of the basic model with heterogeneous regions. The fourth part is devoted to the estimation of the model. In the fifth part we present the application of the model to the analysis of a country stability. In the sixth part we discuss the approach and in the seventh part is the conclusion.

1 Literature review

The core of the literature related to nation formation concerns coalition formation and hedonic game theory. The fundamental paper in this field is Drèze and Greenberg (1980) in which the authors examine the concepts of stability and optimality in hedonic coalitions. A good overview of the existing literature is provided in Hajdukova (2004) paper. One of the recent papers not covered in the 2004 overview, Barberà et al. (2013), presents an analysis of coalition formation under different redistribution rules within the coalitions.

The first political economy papers devoted to the nation formation were published in early

[7]The alternative approach is described in Desmet et al. (2011).

[8]Which can be interpreted as a voting for a particular politician that would represent the region in the parliament and vote for this tax level.

[9]Westhoff (1977).

1990-s and a good overview of the literature of this period is presented in Bolton et al. (1996). In most part of the papers, as Bolton et al. (1996) point out, the heterogeneity in preferences over the public good's provision is modelled using Hotelling location model (Feinstein (1992)). The most broadly known paper that applies this approach is Alesina and Spolaore (1997) that examines the problem of stability and efficiency of the number of the countries and proves that the stable number of countries is not efficient. In this paper the regions' borders are endogenous. Another approach is used in Bolton and Roland (1997). In this paper the authors present a model where the regions are exogenous and consider the impact of factor mobility on the stability of nations.

The literature concerning nations' formation and breakup is not limited only by theoretical papers. In Alesina et al. (2004) the authors provide the empirical analysis of the factors that affect the number of school districts in the U.S. counties based on a an extension of the model from Alesina and Spolaore (1997) paper. In Desmet et al. (2011) the quantitative analysis of the breakup of Yugoslavia is provided. The authors consider a country that consists of several homogeneous regions and assign a specific quasi-linear utility function to the agents living in these regions. They estimate the parameters of this function on real data and they predict correctly the breakup of Yugoslavia in 1990-s. This paper proves that based on existing theoretical models it is possible to analyse quantitatively the stability of countries.

In our paper we follow the logic of Desmet et al. (2011) and expand it for heterogeneous communities. The approach we use is also close to Bolton and Roland (1997) where exogenous borders are imposed. The model we construct in our paper, as well as the one from Desmet et al. (2011), can be used for the analysis of the stability of the large federal states.

2 The model

The main goal of the paper is to construct a model of the large federal state breakup. In theory, that could be modelled as a coalition formation game where each individual is considered as a separate agent (as it is done for the continuum of players in Alesina and Spolaore (1997)). However, as Olson (2009) points out, the equilibria of such games might be impossible to compute within some reasonable time limit, and given that in real states millions of people live, it is definitely the case. As we wanted to construct a model that potentially could be applied to analyse the stability of real countries, we should impose many restrictions to make the model computable.

First, we do not consider each citizen as an agent. We consider a country that consists of N regions. These regions are actually the players of the coalition formation game and the role of the citizens that are assumed to belong to these regions (as we do not allow any migration)

is only in the decision making described below. This makes the approach we use similar to the one applied in Desmet et al. (2011). Next, as we are modelling the possible breakup of the country, we initially assume the existence of a country that should not be necessarily stable (the definition of stability is given below). This approach is common for all the extensions of the model we describe below, but the other assumptions differ.

In the basic model each region of the country is culturally[10] homogeneous but the agents living there are heterogeneous in income. Income cumulative distribution function of region r is $F_r(y)$.

Agents have identical utility functions. The utility depends on the country c and on the region r where the agent lives, on her consumption of private goods x, on the amount of public goods provided by the government g: $U = U(c, r, x, g)$. The agent maximizes her utility function with respect to the following budget constraint: $x \leq (1 - \tau)y$, in which x is the consumption of the private good, τ is the proportional income tax rate and y is the agent's income. This constraint is binding under the assumption that the utility function is increasing in c, we can substitute x to the utility function and get:

$$U = U(c, r, y, g, \tau) \tag{2.1}$$

The utility of the agent depends on the region and the country in which she lives in because this determines the losses from information transmission imperfections and from the cultural differences in public good's provision. We assume that the system of the central government in the state is built such that all the ethnicities living in the country are presented there proportionally to their size. Then, the higher is the cultural distance between the region's culture and the country one, the greater are the losses from heterogeneity. Also, the longer is the distance between the region and the capital of the country, the less efficiently the public good is provided. The motivation for this, as we have explained above, is the imperfection of information transmission.

The government collects $R = \tau Y$ of tax revenues. $G = G(n)$ is spent on the government itself and the rest $R - G$ is provided in the form of a public good. G is a cost function such that $G(n + 1) > G(n)$ and $G(n + 1) - G(n) < G(n) - G(n - 1)$ and we assume it to be continuous and differentiable for simplicity.

The amount of public good per capita is then:

$$g = R - G = Y - G(n) \tag{2.2}$$

[10]Cultural heterogeneity may be replaced by any type of heterogeneity (ethnic, linguistic) other that the income and the geographic ones.

The agents solve the utility maximization problem choosing the optimal tax rate[11]. Next, they choose by majority voting the representative of the region in the central parliament. The representative would be elected if and only if she would vote in the parliament for the tax rate that is chosen through majority voting in the region. Each region has a representative in the parliament. The majority voting for the tax level then occurs in the parliament.

We assume such form of utility function (for example, quasi-linear utility from the example below) that the preferences on the tax rate are single-peaked. Then, from the median voter theorem (Black , 1948) follows that the median tax rate that corresponds in our model to the optimal tax rate of an agent with median income would be the *Condorcet* winner[12].

In the model we allow the unilateral secession (Desmet et al. , 2011) of a region that would occur if the population of the region votes for the separation from the rest of the country. The population makes this decision in a referendum that is organized after every change in the composition of the country (after the secession of some of the regions). We do not allow coalitions of regions to secede and we assume that after the secession the regions become independent states.

One more crucial assumption we make is that the region can not rejoin the country once it has seceded. This guarantees the existence of the stable partition in the model, we would prove this below. For simplicity, we assume that the region with capital is not allowed to secede but this assumption is not necessary[13].

The problem of the agent who lives in region c and has income y is the following:

$$\max_{\tau} \quad U(c, r, y, \tau) \tag{2.3}$$

We assume that the utility function is concave, so the optimal tax rate would be $\tau = \tau(y, r, c)$. As r and c for the agents are the same, the median τ would be the one optimal for the agent with median income $\tau_{me} = \tau(y_{me}, r, c)$. However, notice that the optimal tax rate in the region depends on the geographic distance from the capital and on the cultural distance from the population of the country.

The tax rate for all the country elected in the parliament is $\bar{\tau} = \underset{r}{me}\left[\tau(y^c_{me}, r, c)\right]$ as, again, the preferences of the regions are single-peaked.

The decision about secession is made, as it was claimed above, by majority voting. Given $\bar{\tau}$, the decision to secede is again made by the region's median voter (single peaked preferences

[11] We assume that the agents vote sincerely for their optimal tax rate and do not try to change the outcome of the elections voting for higher or lower rates.

[12] Optimal tax rates for the agents are on R_1, agents are rational and have single - peaked preferences, thus the *Condorcet* winner exists and it is the median tax rate.

[13] This assumption is required for the simulations because there we define country as the coalition of regions that includes the region with the capital.

and one-dimensional choice). The region r secedes if $V(c, r, y_{me}, \bar{\tau}) < V(c', c', y_{me}, \tau')$, c' denotes the single-region country that would be formed by the region that secedes. Note that τ' on the RHS is not the same that was elected as an optimal regional one: τ' is an optimal tax level for a region if the region becomes an independent state.

The obvious gain from secession is the one in cultural and geographic distances i.e. in the efficiency of public good's provision. No imperfect information problem in public good's provision is supposed to arise in a small country with single culture. But at the same time the region would have higher per capita expenses for the government.

The concept of stability we apply is the **unilateral secession stability** concept from Desmet et al. (2011), according to which the country is stable if there is no region in which the majority would vote for the secession. Also this concept is similar to the *laissez-faire* country stability one from Haimanko et al. (2005). From broader point of view this concept is an extension of the *Nash* stable partition as it is defined in Bogomolnaia and Jackson (2002).

It is obvious that under our assumption that the regions can not rejoin the country this stable partition always exists. Assume that there is a country that at the initial point consists of N regions. Then, $k = 0, ..., N-1$ regions could vote for secession. If 0 regions vote for secession, then this is a stable country by definition. If $N-1$ regions vote for secession, in the next round we would have $N = 1$ which is also stable by definition. If $0 < k < N-1$ then in the next round we can assign $N := N - k$ and repeat the procedure above.

Below we present a simple example that illustrates how the model works.

2.1 Example

Consider the case of 3 culturally homogeneous regions in which different cultures live and for which $H_C < H_A < H_B$ and $L_C = 1$; $0 < H \leq 1$ is the cultural distance to the country population and $0 < L \leq 1$ is the geographic distance ($H = 1$ if the country is homogeneous and $L = 1$ if the capital is in this region). Also assume that income distribution is identical for these three regions[14]. Suppose that the individual living in region c has the following quasi-linear utility function (similar to the one used in Desmet et al. (2011)):

$$U = x + \alpha(gHL)^{\beta} \tag{2.4}$$

We do not use additional parameters for H and L in the utility function because at this stage we are interested only in the presence of these factors. However, we can further assume that $H = H(\gamma)$ and $L = L(\mu)$ and analyse the effect of the new parameters on the outcomes of the

[14]To assume that the median income coincides for the citizens of these two regions is sufficient.

model[15].

Tax rate

The agent maximizes the following utility function:

$$U = (1-\tau)y + \alpha((\tau Y - G(3))HL)^{\beta} \tag{2.5}$$

F.O.C. :

$$-y + \alpha\beta(YHL) \cdot ((\tau Y - G(3))HL)^{\beta-1} \tag{2.6}$$

From this follows that on regional level the *Condorcet* winner would be:

$$\tau_r = \frac{G(3)}{Y} + \left[\frac{\alpha\beta}{y_{me}}\right]^{\frac{1}{1-\beta}} [(YHL)]^{\frac{\beta}{1-\beta}} \tag{2.7}$$

From 2.7 follows that, if the regions are almost at the same distance from the capital (and $\beta < 1$), the results are identical to the ones from Desmet et al. (2011) in which the higher is the level of heterogeneity (the lower is H), the lower is the preferred tax rate. In our model the geographic factor affects the decision in the same direction: the longer is the distance between the region and the capital (the lower is L), the lower is the preferred tax rate. Thus, the regions that are situated far from the center, prefer not to finance the provision of public good. This result is reasonable – if the public good is not provided efficiently, then it is not profitable to pay high taxes to finance it. Thus, for such a region the incentives to secede should be higher than for a region which is close to the center *ceteris paribus*.

At the next step on the federal level the representatives would also vote for the median τ. Which region's τ_r would be the *Condorcet* winner depends, from 2.7, only on the relation $\frac{(HL)^{\beta}}{y_m}$. If we assume that the regions have equal median income, then this relation would depend only on HL. The region with median HL would get the tax rate preferred by its median voter as a result of the voting on the federal level. Assume that it is region A that has the preferred tax rate between regions C with the lowest one and region B with the highest.

Then, denote the elected tax rate in the economy as τ_A. The value of the utility function for the agents would be $V = (1-\tau_A)y + \alpha(\tau_A Y - G(3))HL)^{\beta}$.

[15] $0 \leq h$ is some heterogeneity indicator such that the higher is the heterogeneity, the higher is the value of h (similar to the one used in Desmet et al. (2011)). l is a variable with the same properties for the geographic distances.

Secession

In every period the regions vote for the secession after the tax rate is elected. For every region (without loss of generality we write it for region B) for each citizen of the region the decision about the secession would be the choice between $V_{federation} = (1-\tau_A)y+\alpha((\tau_A Y - G(3))H_B L_B)^\beta$ and $V_{unitary} = (1-\tau'_B)y + \alpha((\tau'_B Y_B - G(1)))^\beta$. The latter is the utility in case of secession. L and H equal one and τ'_B is elected. Notice that τ'_B is not the same that we derived solving problem 2.7. It would be τ from the maximization of the following utility function:

$$U = (1-\tau)y_{B,me} + \alpha(\tau Y_B - G(1))^\beta \tag{2.8}$$

Then, the decision of the region is the following:

$$d = \begin{cases} \text{to secede (1)}, & \text{if } V_{unitary} > V_{federation} \\ \text{not to secede (0)}, & \text{if } V_{unitary} \leq V_{federation} \end{cases} \tag{2.9}$$

The stability of country C' under unilateral secession principle is achieved when $d_r = 0 \quad \forall r \in C'$.

We can make a simplest "and never come back" assumption that we mentioned before: the region that seceded from the country in the past is not allowed to rejoin it. We do not consider the regions after secession so we are not interested in whether they unite with other regions to form new countries.

Government expenses effect

One of the effects that we can demonstrate using the example with particular utility function is the effect of government expenses. We have initially assumed that there is a function $G = G(n)$ that denotes the cost of the government. The natural interpretation of this assumption is that the government has some obligatory expenses for national defence, infrastructure and administration without which the state would not be able to exist at all. According to our initial assumption, the agents do not receive any utility from this part of the public goods, and it is more general than to assume that they do.

In the simulations below we show that the volume of these expenses affects the stability of the country. It is the pure effect of the economy on scale.

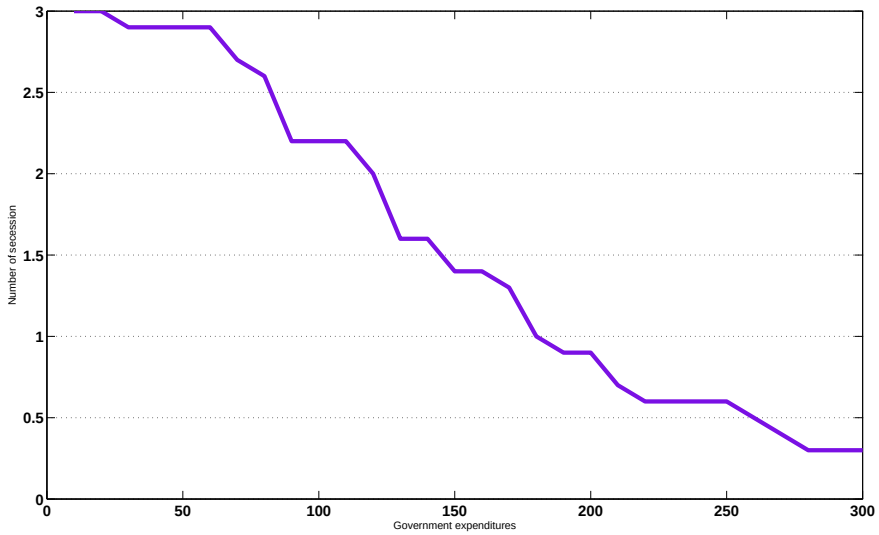


Figure 2.1: The effect of the government expenses on the stability of the countg

We provide an experiment in which we simulate a country with 4 regions. Then, we increase the volume of government expenses and we track the change in the number of regions willing to secede. More precise description of the simulations is presented in the Appendix A.

Figure 2.1 demonstrates the average number of regions corresponding to each government expenditures level that we obtained in our simulations. The result is intuitive – the higher is the volume of the government expenses, the lower is the number of regions that are willing to secede. As the simulations were provided *ceteris paribus*, this picture demonstrates pure effect of the economy on scale on the stability of the country.

Distances effect

We also can show the effect of the geographic distances on the stability of the country. According to the previous argumentation we provided, the geographic distances that are the proxy for the information transmission inefficiency, should be a factor that negatively affects the stability.

We simulated the country breakup process changing the variance of geographic distribution. The simulation proved our intuition – the higher are the distances, the more regions make a decision to secede (Figure 2.2).

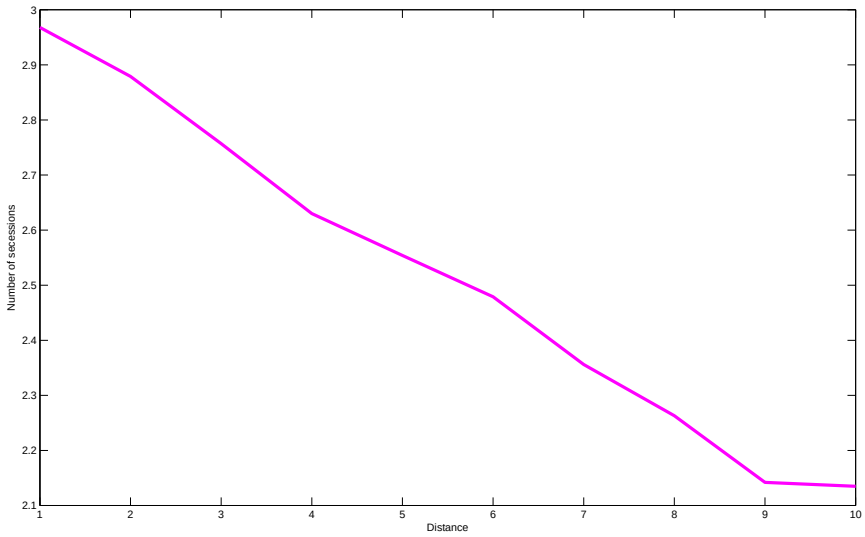


Figure 2.2: The effect of the distances on the stability of the country (higher x means lower distances

The picture would be the same if we change the parameters of heterogeneity instead of geographic distance. Thus, the model proved the intuition that higher heterogeneity and larger geographic distances correspond to lower stability of the country.

3 Heterogeneous regions

In the simplest model above we assumed that the regions are culturally homogeneous. It is obvious that this assumption is rather unrealistic especially in the modern world. In this part of the paper we present a model with culturally heterogeneous regions.

In general, the heterogeneity of population within regions engenders one main difference for the population comparing to the case of homogeneous regions. Now, after the secession from the other parts of the country, agents do not get zero loss from heterogeneity as they now have this loss because of the other cultures in the same region. Moreover, it is possible that after the secession the loss from heterogeneity for some parts of the population would increase (in case if their proportion in the federal state was larger than in the region).

Before starting to analyse the effect of within region heterogeneity, we need to make some simplifying assumptions. We assume that the population of every region is divided in several communities (according to their culture or some other characteristics), each of which has representatives in a regional parliament (we would call it *"Veche"* as the ancient regional "par-

liament" in Novgorod). The parliament is formed proportionally according to the shares of the communities in the population.

The mechanism of the voting is basically the same as in the case of culturally homogeneous regions. At the first step of the voting process each community votes for their optimal tax rate that would be presented in *Veche* by the group of the communities' representatives. It is obvious that this tax rate would be optimal for the median voter of the community. Next, in *Veche* a procedure of voting on the optimal regional tax rate would take place.

It is important to notice here that in this case the winner is not as simply defined as in the case of the homogeneous regions. We can think about this voting procedure as the one when each community is presented by only one representative but her vote is weighted according to the size of the community. Thus, if in a region one of the communities has more than 50% of the population, then its median voter's optimal tax rate would be set as the optimal one in the region.

By assuming this voting procedure, we make a very strong assumption, that the community's representatives would vote for the same tax rate. However, in reality the income factor could be more important than the solidarity with the community. A good evidence of this can be the 2009 Knesset elections where the votes of the population with origins in the Soviet Union (which we would consider as a community by the mother tongue) were divided between left party (Kadima), right party (Likud) and other parties[16]. Thus, the community's interests were presented by the politicians with different point of view on economic policy which contradicts our assumptions.

However, for the empirical analysis based on this model we need number of simplifying assumptions. Namely, we need to determine which agent is the decision maker, which would let us, in empirical studies, collect data only on some particular group of the population.

To demonstrate the major differences between the basic model and this extension, we would assume that the size of the communities is almost equal and thus the voting in *Veche* can be considered as if each community had only one representative. Assuming that each community has the same income distribution, the winner of the majority voting in *Veche* would be the choice of the ethnicity that has median cultural distance to the country.

In case of the secession, again, the same mechanism of voting as in the basic version of the model is applied. Under these assumptions the community with median cultural distance to the country becomes the *Condorcet* winner in voting for secession[17].

[16] Friedrich Ebert Stiftung foundation's report "Israeli Russian Voting Trends in the 2009 Knesset Elections" by Dr. V. Khanin, Bar-Ilan University

[17] Only in the case of almost equal communities.

Mixing the nations

To show the new effects that are observed in case of heterogeneous regions comparing to the basic model, we would compare two artificial countries. In each of these countries there are ethnicities a, b, c presented in equal proportion that belong to separate cultures a, b, c respectively. There are three regions in both of these countries: A, B, C. The only difference between these two countries is that in country 1 the ethnicities live separately in regions A, B, C respectively. In country 2 the ethnicities are equally presented in all the regions, so each region is a mini-copy of country 1 in terms of ethnic composition.

Let's return to the example from the first part of the paper. Assume that two regions are equidistant from the region with capital C. In country 2 in region B (which is equivalent to region A) the voting would lead to the following set of optimal tax rate for the three communities:
$\tau_j = \frac{G(3)}{Y} + \left[\frac{\alpha\beta}{y_{me}}\right]^{\frac{1}{1-\beta}}[(YH_jL_B)]^{\frac{\beta}{1-\beta}}$, $j = a, b, c$

Let's assume that ethnicities a and b are "symmetric" with respect to c, so the distance from a and b to c is h. Then, $\tau_a = \tau_b$ and the *Condorcet* winner in *Veche* would be τ_a. The same would occur in the parliament and τ_a would be elected (in the capital region C the winner would be $\tau_a = \tau_b = \frac{G(3)}{Y} + \left[\frac{\alpha\beta}{y_{me}}\right]^{\frac{1}{1-\beta}}[(YH_a)]^{\frac{\beta}{1-\beta}}$, but the decision would be made by two identical regions A and B).

If region B secedes from the country, then the tax rate set up on *Veche* would be $\tau'_a = \tau'_b = \frac{G(1)}{Y_B} + \left[\frac{\alpha\beta}{y_{me}}\right]^{\frac{1}{1-\beta}}[(Y_BH_a)]^{\frac{\beta}{1-\beta}}$. Using this information, we can compare the indirect utilities in case of secession and in case of staying inside the state for median agent in the region.

$$V_{federation} = (1-\tau_B)y + \alpha((\tau_B Y - G(3))H_bL_B)^\beta \tag{3.1}$$

$$V_{unitary} = (1-\tau'_b)y + \alpha((\tau'_b Y_B - G(1))H_b)^\beta \tag{3.2}$$

Notice that in both cases the loss from heterogeneity multiplier is H_b.

If we address to country 1 with homogeneous regions then we would get the same $V_{federation} = (1-\tau_B)y + \alpha((\tau_B Y - G(3))H_bL_B)^\beta$ (and the tax rate τ_B would be the same) but $V_{unitary} = (1-\tau'_B)y + \alpha((\tau'_B Y_B - G(1)))^\beta$.

In both countries region B would vote for secession only in case if $V_{federation} - V_{unitary} < 0$. Assume, that the parameters are such that in case of heterogeneous regions this is not true – $V_{federation} - V_{unitary} \geq 0$. We want to prove that in case of homogeneous regions it could be that, given the same ethnicities and the same distances of the regions to the capital, it is possible that $V_{federation} - V_{unitary} < 0$. We have already claimed above that $V_{federation}$ in both cases is

the same. Then, we need only to prove that V is decreasing in heterogeneity.

Returning to the results derived above, ${\tau'}_a = {\tau'}_b = \frac{G(1)}{Y_B} + \left[\frac{\alpha\beta}{y_{me}}\right]^{\frac{1}{1-\beta}} [(Y_B H_a)]^{\frac{\beta}{1-\beta}}$ and $V_{unitary} = (1-\tau'_b)y + \alpha((\tau'_b Y_B - G(1))H_b)^{\beta}$. If we substitute the rate into the expression for $V_{unitary}$, we would get

$$(1 - \frac{G(1)}{Y_B} - \left[\frac{\alpha\beta}{y_{me}}\right]^{\frac{1}{1-\beta}} [(Y_B H_b)]^{\frac{\beta}{1-\beta}})y + \alpha\left((\frac{G(1)}{Y_B} + \left[\frac{\alpha\beta}{y_{me}}\right]^{\frac{1}{1-\beta}} [Y_B H_b]^{\frac{\beta}{1-\beta}} Y_B - G(1))H_b\right)^{\beta} \quad (3.3)$$

Expression 3.3 positively depends on H_b[18], thus, as the higher is the distance, the lower is H_b, we can conclude that $V^1_{unitary} > V^2_{unitary}$ and thus there can be such a situation that in country with heterogeneous regions the break up will not occur but in the one with homogeneous it will.

Does this result hold in general? The answer is no. We have shown the difference *ceteris paribus* – with the same income distribution, even with the same ethnicities living in regions. If the heterogeneity in the region is small comparing to the overall country, then the phenomenon shown above would be observed.

However, what we have shown in this example is that regional heterogeneity potentially could make the country more stable.

We should additionally notice that here we implicitly assumed that the loss of heterogeneity within the region is the same as within the country. In reality it can be very different and the result we derived above would not be relevant.

4 Estimating the model

In the empirical part of paper we use the simplified version of the model presented in previous sections. That is the the model with heterogeneous regions which takes into account both cultural and geographic distances, but we assume that the agents get utility from public good which includes the spendings G. Thus, in terms of our model, $g = R = \tau Y$, where g is the public good, R is the tax revenue and Y is the total income (of the region or of the country).

The reason for this is the difficulty of the identification of G. In our definition G are the government expenditures the agents do not benefit from. It is easy to believe that the population does not benefit from some types of expenditures directly (for example, a median citizen does not get any utility from the fact that the army gets new jet-fighters if there is no war, at least), but to divide the total expenditures into two groups does not seem to be a possible task, at least, it requires additional analysis which includes the collection of data on population's perception

[18] For some values of the parameters, not generally.

of government spendings. Instead, we assume that the agents receive the utility from all the public goods provided by the government.

Thus, we need to estimate every parameter of the utility function

$$U = (1 - \tau)y + \alpha(\tau Y H(\gamma_h, h)L(\gamma_l, l))^{\beta} \tag{4.1}$$

H is the measure for heterogeneity, L is a measure of the geographic distance from the region to the country, that depend on the corresponding distances h and l and parameters γ_h and γ_l. These parameters have simple interpretation – they determine the effect of cultural distances and geographic isolation on the effect of the public good provision.

The parameters of the utility function are not simultaneously identifiable because the number of regressors in the resulting econometric model is lower than the number of the parameters. For this reason we use technique other than standard econometrics tools. Two parameters (α and β) can be identified if we consider only small countries with homogeneous population. This was done in Desmet et al. (2011) and, as on this step we need to replicate completely these estimates, we just use this paper's result. The Table 1 presents the estimates of the coefficients we use. These estimates were calculated using different samples and only the first one satisfies the small-and-homogeneous country assumptions.

	α	β
Only linguistically homogeneous countries	22.05	0.095
All countries	25.80	0.0833
All countries ($H = 0.3$)	51.05	0.052

Table 1: Estimates of α and β. Source: Desmet et al. (2011)

The second part of the empirical work consists of the estimation of γ_l and γ_h – the parameters of the measures of heterogeneity and geographic distance. As we mentioned before, we are not able to identify these parameters due to the lack of regressors. For this reason we use the technique proposed in Desmet et al. (2011) and we find the set of possible values of these parameters.

Namely, we claim that these parameters should satisfy the conditions under which the current partition of countries is stable. Thus, we consider two sets of countries. The first set consists of countries in which there are unstable regions[19] (e.g. Spain and the Basque country) and we find such γ_l and γ_h that the secessions do not happen. The second set consists of countries that

[19]By this we mean regions in which there is a significant separatist movement

already broke up and we do the same thing: we are looking for parameters for which the break up should have happened.

The complication of this approach is that the optimal tax rate of the regions that have not seceded is not observable (Basque country, Flanders). Thus, for each comparison we have to construct a measure for the optimal tax rate using our theoretical results (Equation 4.2).

Namely, we would compare:

United country

$$\mathrm{V}(\text{united}, \tau_{observed}) \geq \mathrm{V}(\text{separated}, \tau_{unobserved})$$

Separated country

$$\mathrm{V}(\text{united}, \tau_{historical}) \leq \mathrm{V}(\text{separated}, \tau_{observed})$$

By $\tau_{historical}$ we mean the tax rate set in the regions before the secession. It is observed, so generally speaking we compare the utility just before the secession and after the secession.

We construct the following measure for $\tau_{unobserved}$ ($L = 1$ because the capital is now inside the region):

$$\tau_{unobserved} = \left(\frac{\alpha\beta}{y}\right)^{\left(\frac{1}{1-\beta}\right)} (YH)^{\left(\frac{\beta}{1-\beta}\right)} \tag{4.2}$$

where Y is the total income of the region, H is the heterogeneity measure and y is the median income in the region. The important thing to do is to define the decision-making community or the *median* community in the sense that its decision would determine both the secession and the tax rate in the region. It is easy to determine which community is such a decision maker in case if there is a community that gets more than 50% of votes in the regional parliament. It is more difficult in case if the region is divided into multiple relatively equal in size communities (e.g. Dagestan republic in Russia), then we need to calculate the optimal tax rates for all the communities as well as the secession decisions for all the communities. The reason for this is that the optimal tax rate is strictly increasing in H (see 2.7) but the decision on secession is not (see, for instance, 3.3).

From each of these comparisons we get set $\Gamma_i : (\gamma_h, \gamma_l) \in \Gamma_i$. The feasible set would be $\Gamma = \bigcap_i \Gamma_i$. That is the final step of the estimation procedure.

Concerning the application of the estimated model, we would take the data for the country j and find such $\bar{\Gamma}_j$ that secession does not occur and then compute the ratio of the areas of $\bar{\Gamma}_j$ and Γ which would be the probability that the state is stable and the secession would occur. By doing this, we assume the uniform distribution of the parameters over the set.

Technical implementation

The corner stone of our analysis is the community voting system. We define communities according to the mother tongues and then use the linguistic distances as proxies for more general cultural distances. We assume the following form of linguistic and geographic distance measures:

$$L = exp(-\gamma_l l)$$

and

$$H = exp(-\gamma_h (h-1))$$

We start with 402201 points of the grid $\gamma_l \in [0;2]$ with step 10^{-3} and $\gamma_h \in [0;0.02]$ with step 10^{-4}. Our results presented below suggest that this initial grid is sufficient for the calibration of the model.

The first cases on which we calibrate the model are the Basque country and Alsace. We chose these two cases on purpose – in the Basque country there exists a strong separatist movement and Alsace is a region with a complicated history on the border of France and Germany. Both regions are situated relatively far from the capital and in both there are large shares of population with mother tongue different from the official language of the country. On Figure 4.1 the separate Γ_i are shown.

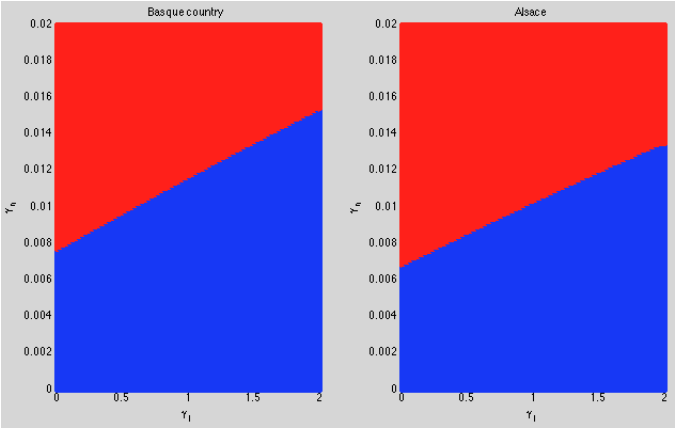


Figure 4.1: Initial set adjusted after adding the Basque Country and Alsace data. Blue area is the set of admissible values of the parameters.

Technically, the procedure is the following (consider the Basque country case):

1. Collect data on median income, total income, linguistic and geographic distances for the country and for the region. At this step we consider only major communities not taking

into account the small immigrant ones.

2. Define the decision-making community in the region (native Spanish speakers of Basque Country in our case). As it was said before, in cases that we consider it is an easy task to define the decision-making community – it is the largest community in the region. However, in cases when the regions are divided into multiple small communities neither of which has the majority in the regional parliament, decision for each of the communities should be computed and then the median decision should be calculated using the weighted voting described in the theoretical part of the paper.

3. Collect data on share of tax income in GDP.

4. Estimate the optimal tax rate in the region in case of secession using equation 4.2.

5. Calculate the utility for the region in case when it is part of the country and in case if it secedes for each point of the initial grid defined above.

6. For each element of the initial grid compare the values of the utility function for these two cases.

7. Exclude from the set points in which the regions should have seceded.

The adjustment of the grid that we provided above let us find the upper bound of the values of the parameters. The next part of the calibration is the adjustment of the grid on real breakup cases which would let us obtain the lower bound. The model we suggest in our paper does not consider possible unification of several regions (or countries) to one, thus we can not obtain the lower bound using conditions under which modern countries would stay independent (like it was done in Desmet et al. (2011)). Instead, we consider *post factum* the cases of breakup looking for the parameters for which the secession was a right decision for the country.

There were several breakup cases in the last decades of the XX-th century: Yugoslavia, Czechoslovakia and the Soviet Union. We would calibrate the model on the clash of the Soviet Union[20]. Potentially, if we consider Russia as the successor of the Soviet Union, we can calibrate the model using 14 cases of the other republics. However, we would exclude from the calibration the cases of Kazakhstan and Belarus, because they became the parts of the Eurasian Economic Community and because Russia and Belarus are parts of the Union State. Also, we exclude Ukraine from the calibration due to recent events.

[20]Due to the availability of the data.

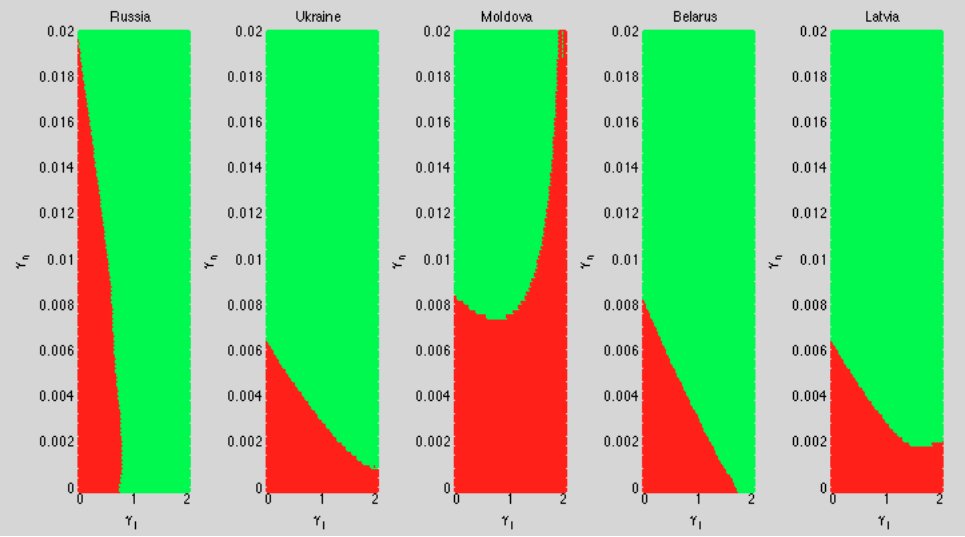


Figure 4.2: The calibration on the clash of the USSR case.

The best strategy of the calibration is to obtain the data for the last several years of the existence of the Soviet Union and for several years after its clash and to compare directly the utilities for each country thus calculating the bounds of the parameters under which the breakup would make countries better off. However, due to disorder of the period of late 1980-s and early 1990-s and to the absence of the comparable data for the rest of the period till now, we use the data for 2011 and construct the measure for the optimal tax in an artificial Soviet Union of this period of time using our model. First, for each of the countries we find the decision-making community. Next, we calculate the optimal tax rate and find the median across the Soviet Union which should be set according to the community voting system. Finally, we compare the utility in case of a unified state and in case of independent states. As we expected, this allows us to obtain the lower bound for the parameters (Figure 4.2 and Figure .1).

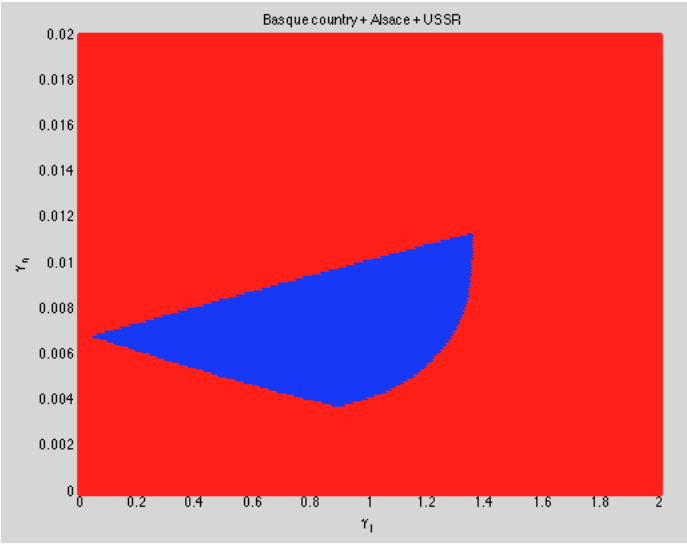


Figure 4.3: The final set

We used data for Spain (Basque country cases), France (Alsace) and USSR (described below) to obtain the grid which is on Figure 4.3. The summary statistics is presented in the Appendix D.

Thus, we obtain the set Γ assuming that Alsace and the Basque country are stable and the USSR republics are better off as independent states than they are as parts of the Soviet Union (Figure 4.3). Technically, we started with 402201 points of the grid $\gamma_l \in [0; 2]$ with step 10^{-3} and $\gamma_h \in [0; 0.02]$ with step 10^{-4} and the adjusted grid consists of 51390 points. We should note that the adjusted area is included into the set from which we started thus our choice of the initial grid was relevant and expanding the initial grid would only increase the computational time of the adjustment of the model.

5 Application

We calibrated the model on several cases (Alsace, the Basque country and several USSR republics). By doing this, we obtained a set of values for which the current partition is optimal is the sense of decision-making agents' utility. Thus, the further investigation and the application of our results are based on the assumption that Alsace and the Basque country are stable regions and that the Soviet republics are better off being independent states.

We apply the model to the cases of Quebec, Catalonia and Flanders. These three regions satisfy the assumptions under which our model can be applied – primarily, they have relatively high degree of autonomy. The case of Catalonia is especially interesting, because we analyse its stability given that another Spanish region, the Basque country, is stable.

Technically, the procedure is the following. In our previous notations, we find $\bar{\Gamma}_j$ which is the

intersection of Γ and the set of parameters for which the region under consideration is stable. Then, we compute the probability $p_j = 1 - \frac{\#\bar{\Gamma}_j}{\#\Gamma}$ of secession[21].

Region	p_J
Flanders	1
Quebec	1
Burgundy	0.1017
Catalonia	0.9875
Russia	0.7089
Belarus	0.9

Table 2: The estimated probability of breakup.

The results suggest that under the assumptions we made Quebec and Flanders should secede and Catalonia would also secede with probability close to 1. Let's consider each of these cases. The population of Quebec is mostly French - speaking while in Canada the majority are English - speaking[22], but the region benefits from being a part of Canada in terms of public goods – median income there is lower than for all the country (Table 3) and the regional total income is approximately 20% of the country one[23]. Thus, Quebec benefits from the independence which results in the possibility to set the tax rate and to improve the quality of the information transmission more than from higher amount of public good. In case of Flanders, Dutch-speaking part of Belgium, the situation is similar and Flanders does not lose as much as Quebec in terms of public goods as the share of this region in Belgian economy is larger[24]. Both in Quebec and in Flanders the decision making communities were different from the majority of population of the rest of the country (French-speaking and Dutch-speaking communities respectively). The case of Catalonia is different as the majority of population has Spanish mother tongue. Catalonia, a relatively rich North-Eastern region of Spain, benefits from the increase of the information transmission efficiency and from the possibility to set the tax rate. At the same time the decision-making community is worse off in terms of heterogeneity and of the public good provision in case of secession. The values of the parameters for which the regions are stable are presented on Figure 5.1.

[21] # means the number of the elements of the grid. In theory p should be equal to the relation of the areas of these two sets.

[22] According to 2011 Canada Census

[23] OECD Regional Statistics Database

[24] OECD Regional Statistics Database

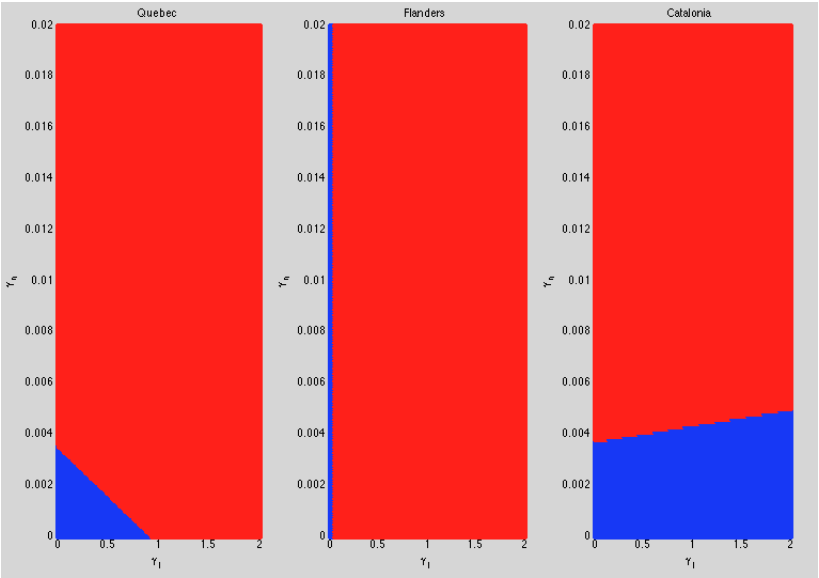


Figure 5.1: Quebec, Flanders and Catalonia. Blue dots – stable.

We considered quite unstable regions and obtained probabilities close to one or equal to one. We should also demonstrate that the model works in less extreme cases. For this purpose we considered the cases of Belarus and Russia secession from the Soviet Union. The results suggest that the probability that Russia secedes is only 0.71 (Table 2). In addition, we applied the model to a region that we consider as stable – French Burgundy. As expected, the probability of secession is very low – around 0.1.

Thus, we applied the model to regions that are considered as unstable, to a stable region and to the countries that already became independent. In all three cases the estimated probability of secession was adequate and consistent with our expectations. Thus, we can conclude that this model can be applied to the analysis of the stability of the countries. However, the results should be treated very carefully – they suggest that assuming that the Basque country and Alsace are stable and that the Soviet republics are better off being independent states, the model provides a particular probability of secession. The results are not robust to changes in the set of countries on which the model is calibrated, thus they should be treated only conditionally on this calibration set.

6 Discussion

The unilateral secession concept we used could sometimes be irrelevant for the modelling of the countries' breakup because it is unlikely that any part of the federation can harmlessly secede. The alternative to it, also mentioned in Desmet et al. (2011), is the *Limited Right*

of Map Redrawing that requires the agreement on secession from the rest of the country. It is very easy to modify the model we suggest for this concept of stability – we just need to assume that the voting on the secession should be "organized" not only in every part of the country r, but also in the corresponding $\Omega \backslash r$ and only in case if the decision to secede is approved in both voting procedures, it should occur. It can change empirical results a lot – two elections instead of one should make countries more stable.

Another reason for critique is that in our model we do not assume that regions are subject to some "fine" for the secession as it is described in Bolton and Roland (1997). Again, this would be only an additional complication of the model but it would not change the results (but would obviously make the country more stable). Obviously, regions of the countries are held together not only by the public goods but also by trade, administrative, transport connections. For this reason we can not estimate the potential losses from secession.

Our paper could also be criticized for the Nash equilibrium concept. By considering only this concept of stability, we do examine possible secessions of coalitions of players. We believe that this approach is relevant because history does not know many examples when country's parts seceded as a coalition. For example, the Baltic countries ,that were part of the Soviet Union, seceded as independent states. The Basque country, Catalonia, Scotland and other regions that have separatist movements also do not have intention to secede as coalitions with other regions. Thus, we can assume based on these examples that our model in which regions make the secession decision separately is relevant.

We should note that we consider a model which describes only unilateral secessions under assumption that the region becomes independent and does not join another country. The land partitions that are the results of the wars of some governments' deals are not described by this model, neither are the cases when a region secedes to join other country while the latter can be described by obvious extension (in this case, the utility value for the secession should be calculated as a value in case if the country becomes part of the other country).

Also we should stress the fact that the tool for the analysis of the countries' stability we provide here should not be used for the justification of any political decisions. The model we suggest is a very rough approximation of the processes that take place, many elements are missing (e.g. transfer structure and trade patterns) and its only application is to show the potential instability in a region. Also, as it was mentioned above, the resulting probabilities of secessions should be treated conditional on the initial calibration set of countries.

7 Conclusion

The state formation is the process in which a country should accept the bride (the increase in heterogeneity) to get the dowry (the economy on scale in public good's provision). The nation breakup is the reversed process in which parts of the country can gain from their independence by giving up the economy on scale benefits.

In our paper we presented a model of a large federal state breakup process. As we can see from the simple example provided above, the geographic distance affects the agents' decisions in the same way as cultural distances do. However, it is important to distinguish between the influence of these two factors because cultural distances are not perfectly correlated with the geographic ones and there could be regions close geographically but not culturally and *vice versa* . The introduction of the geographic distances as a proxy for the information transmission imperfections, even in these simple settings, contributes to the understanding of the breakup processes.

The main contribution of this paper is the construction of the model with heterogeneous regions which can be used as a tool for the analysis of the stability of the countries. We suggest a way to deal with complications in the empirical analysis caused by the heterogeneity of regions' population. We calibrate the model assuming that the Basque country and Alsace are stable regions and we apply the model to the analysis of the stability of Catalonia, Quebec and Burgundy.

Our results should be treated very carefully. It is important to note that these results are based on multiple assumptions, such as the form of the utility function, the unilateral secession principle, the specific set of stable regions, and they should not be used to justify any political actions.

References

Alesina A., Baqir R., Hoxby C., Political Jurisdictions in Heterogeneous Communities *Journal of Political Economy* ,2004, Vol. 112, No. 2, pp. 384-396

Alesina A., and Spolaore E., On the Number and Size of Nations. *The Quarterly Journal of Economics* , Vol. 112, No. 4 (Nov., 1997), pp. 1027-1056

Ashraf Q., Galor O. and Özak O. Isolation and Development *Journal of the European Economic Association*, April–May 2010 8(2–3):401–412

Salvador Barberà, Carmen Beviá, Clara Ponsatì, Meritocracy, Egalitarianism and the Stability of Majoritarian Organizations. *ISER Seminar Series* 4 10 13

Black D., On the Rationale of Group Decision-making, *Journal of Political Economy* Vol. 56, No. 1 (Feb., 1948), pp. 23-34

Bogomolnaia, A., Jackson, M. O. , The Stability of Hedonic Coalition Structures, *Games and Economic Behavior* 38 (2002), p. 201-230.

Bolton P., Roland G., Spolaore E., Economic theories of the break-up and integration of nations. *European Economic Review* 40 (1996) pp. 697-705

Bolton P., Roland G., The Breakup of Nations: A Political Economy Analysis. *The Quarterly Journal of Economics* , Vol. 112, No. 4 (Nov., 1997) pp. 1057-1090

Desmet K., Ortuño-Ortín and Weber S., Peripheral Diversity and Redistribution, *CEPR Discussion Paper* No. 5112 (June 2005)

Desmet, K., Breton, M., Ortuño-Ortín, I., Weber, S. The stability and breakup of nations: a quantitative analysis. *Journal of Economic Growth* Vol. 16, Issue 3., pp. 183-213

J. H. Drèze and J. Greenberg, Hedonic Coalitions: Optimality and Stability. *Econometrica*, Vol.48, No. 4 (May, 1980), pp. 987-1003

Fearon J.D. and Laitin D.D., Weak states, rough terrain, and large-scale ethnic violence since 1945 *Paper prepared for delivery at the 1999 Annual Meetings of the American Political Science Association, 2-5 September 1999, Atlanta, GA*

Feinstein, J.S., Public goods provision and political stability in Europe, *The American Economic Review*, Vol. 82, No. 2, Papers and Proceedings of the Hundred and Fourth Annual Meeting of the American Economic Association (May, 1992), pp. 323-329

Haimanko O., Le Breton M., Weber S., Transfers in a polarized country: bridging the gap between efficiency and stability. *Journal of Public Economics*, 89 (2005), pp. 1277–1303

Hajdukova, J., On coalition formation games. *Technical Report, Institute of Mathematics, PJ Safarik University* (2004)

Kyriacou A., Palacín N.M., Secessionism and the quality of government: evidence from a sample of OECD countries *Working paper* (2014)

Olson M., Nash Stability in Additively Separable Hedonic Games and Community Structures, *Theory of Computing Systems* (2009) 45: 917–925

Westhoff, Existence of Equilibria in Economies with a Local Public Good, *Journal of Economic Theory* 14, 84-112 (1997)

Appendix A

We assume that the agents' income is distributed uniformly on $[1, 3]$. The initial state consists of 4 regions each of which has 1000 population. We generate heterogeneity and distance multipliers as the numbers uniformly distributed on $[0.7, 1]$. For α and β we use parameters close to the ones from Desmet et al. (2011): $\alpha = 20$ and $\beta = 0.05$. Then, we simulate the economy 100 times and for each case we find the number of regions willing to secede given G. Here, for simplicity, we assumed that G does not depend on n. The average for 100 simulations is presented on Figure 2.1.

Appendix B

Making the same assumptions as in the first simulation (4-region country with 1000 population in each region, uniformly distributed income on $[1, 3]$), we gradually change the variance of the initial distribution of points on the line given $G_0 = 100$. We start with uniform distribution on $[0, 1]$ (which gives potential maximum distance to the capital 1), next $[0, 0.5]$, $[0, 0.33]$ and the last is $[0, 0.1]$.

Appendix C

We would construct a simple model that provides the intuition about the relation between the distance from the region to the center of the country and the efficiency of public good's provision.

Assume that a public good G provided in the region is divided into several "projects" $G_1, G_2, ..., G_n$. The realization of these projects is the responsibility of the local authorities.

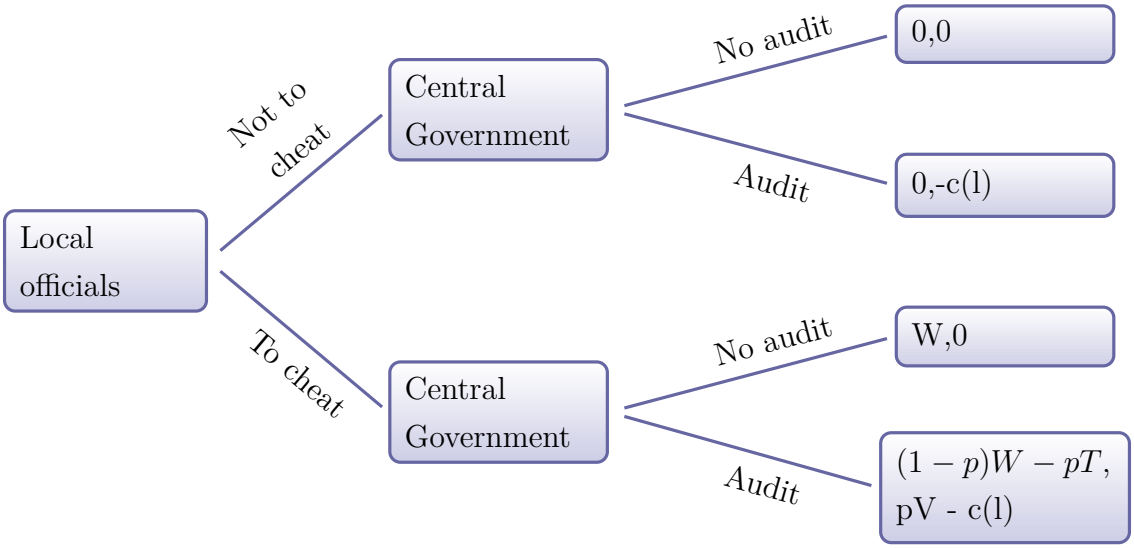
The local authorities can provide the projects inefficiently and cheat (we do not specify how) or they can provide the projects efficiently. The central government can either provide audit or not. The game is repeated for every project.

Cheating would give to the local authorities the payoff W. In case if the cheating, the local authorities would lose T.

The cost of the audit is $c = c(l)$ – an increasing in distance l from the center of the country function. The probability to reveal cheating when the audit is provided is p. For revealing the cheating the central government gets V. On the diagram below we show the expected payoffs of the players.

We can show that if $pV - c(l) < 0$ and $(1-p)W - pT > 0$ than the Nash Equilibrium would be $\{\text{To Cheat}, \text{No audit}\}$. Thus, if the price of the audit is high (which from the assumption

would happen when the distance is large), the local authorities would cheat in the equilibrium and the public good would be provided inefficiently.



The game between the Central government and the local officials

Appendix D

Region	τ	$y_{med,c}$	$y_{med,r}$	h_c	h_r
Basque country	.325	19288	25076	1.51	3.06
Alsace	.429	23696	23521	1.44	3.40
Burgundy	.429	23696	21 541	1.44	1
Flanders	.435	25429	22256	4.08	1
Quebec	.306	24228	21497	3.48	1.81
Catalonia	.325	19288	22686	1.51	1.82

Table 3: Data on median primary income (2010, PPP current dollars) and linguistic distances for the decision-making groups *Source: OECD Regional Statistics Database and Country-level Censuses*

Appendix E

We follow Fearon and Laitin (1999) and use the linguistic trees provided by Ethnologue to calculate the distances between the languages. The languages we deal with have maximum length of the tree (number of divisions to branches) equal 11. We calculate the distances between

languages in the following way : $h = 11 - n$ where n is the level on which the languages "break off" on the tree. For example, the Basque language is an isolate, so it is separated from all other language families, thus it breaks off at $n = 0$ and $h = 11$ for it. For Catalan and Spanish the break off is on the 7-th level and thus $n = 7$ and $h = 4$. The distance between a language and the language itself is 1.

Appendix F

In the choice of stable countries we refer to Kyriacou and Palacín (2014). In this paper measure the secessionism in OECD countries as the vote share of secessionist parties in the election of the national parliament. The authors argue that within the period of observations (1980-2007) the following countries had zero share of secessionist parties in the parliaments: Australia, Austria, Finland, Greece, Hungary,Ireland, Netherlands, New Zealand, Slovak Republic, Sweden and Switzerland. The rest of the OECD countries can be ranked in the following way, using 2004-2007 elections data presented in Kyriacou and Palacín (2014):

Rank	Country
1	Portugal
2	U.S.A.
3	Germany
4	Czech Republic
5	Denmark
6	U.K.
$> 5\%$	
7	Italy
8	Spain
$> 10\%$	
9	Canada
$> 20\%$	
10	Belgium

Table 4: Countries with non-zero share of nationalist parties in the parliaments ranked ascending using Kyriacou and Palacín (2014) data on 2004-2007 elections.

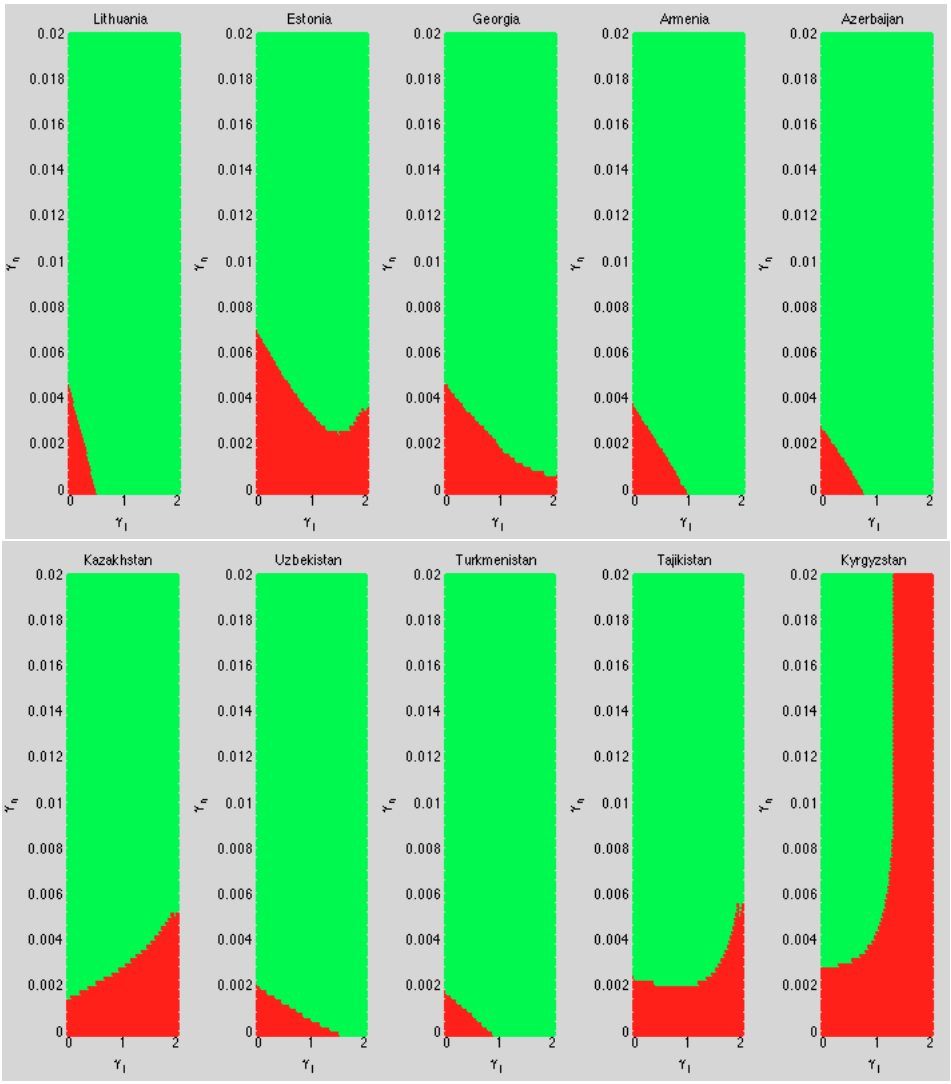


Figure .1: The calibration on the clash of the USSR case.