Universität Trier

National Identity and Religious Diversity

Kenneth Harttgen
Matthias Opfinger

Research Papers in Economics
No. 7/12

National Identity and Religious Diversity

Kenneth Harttgen, Matthias Opfinger *

ETH Zurich, University Trier

This version: 4 December 2012

Abstract

We construct a simple index for national identity using information from the World Values Survey on peoples' attitudes towards the nation. We then analyze the relationship between national identity and religious diversity. The results show that religious diversity is significantly and negatively related to national identity. We also find support for the previous finding that ethnic diversity does not seem to be related to national identity. Apparently, religious diversity and the feeling of heterogeneity that goes along with it prevents people from forming a common identity. Besides, our results indicate that there is a negative correlation between national identity and the level of income. The prevention of discrimination against minority groups follows as an important policy advice.

JEL-classification: J15, O1, Z12

Keywords: National Identity, Religious Diversity, Identity Formation

*ETH Zurich, Nadel, Voltastrasse 24, 8092 Zurich, Switzerland. Phone: +41-(0)44-6329825. E-Mail. harttgen@nadel.ethz.ch. University Trier, Department of Economics. Universitaetsring 15, 54286 Trier, Germany. Phone: +49-(0)651-201 2714. E-Mail: opfinger@uni-trier.de. We thank Carl-Johan Dalgaard, Sebastian Vollmer, two anonymous referees, and conference participants in Forli, Glasgow, and Goettingen for helpful comments.

1 Introduction

Large empires have disintegrated into smaller units over the last centuries or in some cases even only in the last decades. New entities have been formed along specific lines which are today known as nation states. But how did these states evolve? What do the inhabitants have in common? The analysis of nations is popular in the political and social sciences. It is investigated by, e.g. Anderson (2006), Alesina and Spolaore (2003), Bloom (1990), Gillis (1996), Miller (2000), Triandafyllidou (1998), Wodak et al. (1998), or concerning the role for specific nations or Europe by, e.g. Checkel (1999), Maier (1997), Noiriel (1996), or Smith (1992).

Another strand of the literature explores the determinants for social cleavages and civil wars. Collier and Hoeffler (1998), Vanhanen (1999), and Fearon and Laitin (2003) investigate the economic determinants of civil wars. Vanhanen (1999) finds that higher heterogeneity increases the probability of the occurrence of civil wars. Collier and Hoeffler (1998) argue that the effect is not linear. First, the probability of a civil war rises with higher levels of ethnic diversity but after a maximum is reached further increases in ethnic diversity reduce the probability. Thus, Montalvo and Reynal-Querol (2005, 2005a) calculate a measure of ethnic polarization, thereby following an idea of Esteban and Ray (1994). This index reaches its maximum if the society consists of two large rivaling ethnic groups. Montalvo and Reynal-Querol (2005) show that ethnic diversity has a direct negative impact on the GDP growth rate, whereas the impact of ethnic polarization and religious polarization is indirect through reduced investment, increased government consumption, or a higher probability of civil wars. In a following paper (2005a) they analyze the direct impact on civil wars and find that increasing ethnic polarization has a significant positive impact on the occurrence of civil wars.

Apparently, homogeneity plays an important role for the smooth working of a society. An important question that arises is if simple differences in ethnicity or religious beliefs cause social trouble or, as might as well be the case, feelings of commonness reduce rivalry between different groups (Tajfel, 1970). If people feel close to their neighbors due to shared values or beliefs there might be no reason for social cleavages even if the neighbors are ethnically different. Georgiadis and Manning (2009) call this "nation-building", other authors speak of a national identity[1].

If a common identity reduces the probability of civil wars it is important to investigate the relationship between the formation of a national identity and social heterogeneity. Miles and Rochefort (1991), Calhoun (1993), Jones (1997), and Bond (2006) analyze the relationship between ethnic diversity and national

[1] Smith (1991) devotes a book to the question of national identity explaining the creation of nation states and proposing many determinants and channels which might influence the formation of a national identity.

identity and suggest that ethnic diversity and national identity influence each other, lacking, however, a numerical concept of national identity. In a recent study Masella (2011) finds that ethnic heterogeneity does not have a significant effect on national identity.

We contribute to this literature proposing that religious diversity is more important for national identity than ethnic diversity. We generate a measure for national identity building on answers to questions concerning the attitude towards the nation from the World Values Survey. We then analyze the relationship between this index of national identity and religious diversity. Figure 1 shows a scatterplot for national identity and religious diversity on the country level.

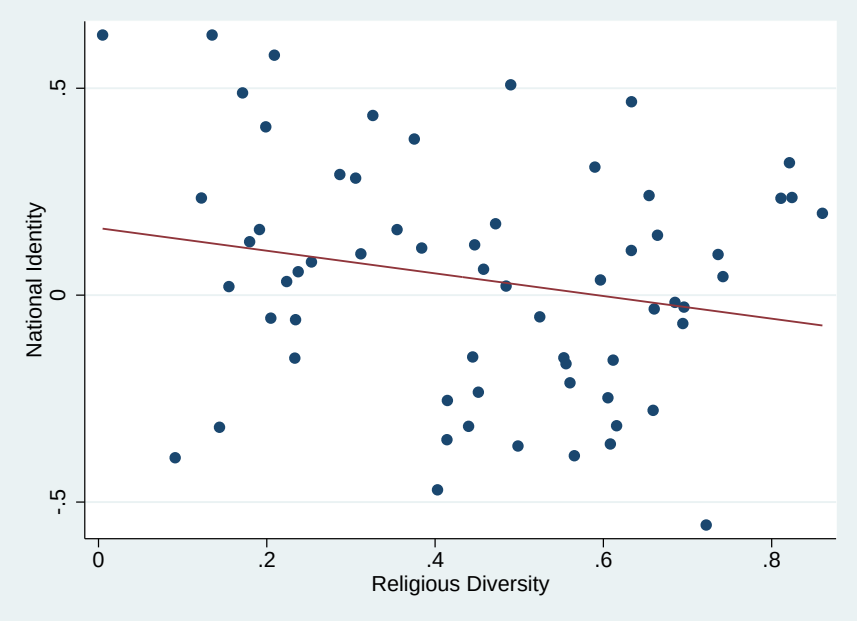


Figure 1: Relationship between Religious Diversity and National Identity

Religious diversity may be more important for the formation of a common identity because the choice of a religion reveals certain preferences and generates a set of common values and beliefs. However, these values can differ substantially between different religious denominations. Furthermore, every human can choose his religion whereas race and ethnicity are determined by the genes. Consequently, we suggest that the active choice of common values is more important for a shared identity than coincidental ethnic differences. Our aim is to analyze the relationship between national identity and religious diversity, where we expect to find a negative relationship. If this is true we will try to explain this relationship taking into account that the relationship with ethnic diversity is not statistically significant.

The concept of identity has been introduced into the economics literature by Akerlof and Kranton's

(2000) influential article. They add identity to the utility function and can thereby explain why some outcomes are optimal for a group of people while they might be detrimental to others. Identity can affect economic outcomes through changes in the payoffs from own actions or from the actions of others. Furthermore, the choice of an identity can affect economic behavior or changing social norms might alter identity-based preferences.

In the following years several studies have been conducted which further investigate the role identity might play for economic outcomes. Bisin et al. (2010), for example, disentangle the identity formation process and propose two mechanisms. Cultural conformity claims that minority groups adopt inclusive identities and that they integrate into their social surroundings. Contrary, cultural distinctiveness proposes that minorities keep their identities and reduce interactions with individuals from other ethnic groups. The authors find empirical evidence supporting the idea of cultural distinctiveness. Darity et al. (2006) provide an evolutionary model that discusses inter- and intraracial interactions based on identities and explains under which circumstances racialist or individualistic identities are formed.

Bodenhorn and Ruebeck (2003) analyze the identity formation process of African Americans in the Antebellum South and find that the size of the community determines the probability of choosing a mixed-race identity. Similarly, Austin-Smith and Fryer (2005) find that the cost of leaving the peer group explains the education decisions of African Americans. By "acting white", i.e. becoming better educated, African Americans lose their former identity and choose to integrate into the white, presumably rich, network. Battu et al. (2007) come to a very similar conclusion when they investigate job market decisions of non-whites. Peer pressure and the possible gains of adopting a white identity heavily influence job market decisions. Constant and Zimmerman (2008) and Constant et al. (2009) develop a measure of ethnic identity and investigate why migrants might choose an identity that favors the country of origin over their host country.

These studies suggest that identity has an important influence on personal economic outcomes. However, we are interested in a national identity and how it is influenced by social heterogeneity. That is why we calculate our measure for national identity on a national basis and than investigate the role of diversity on a country level. In our empirical section we show indeed that high levels of religious diversity within a country are related to lower levels of national identity, thereby confirming the hypothesis that different values might reduce the bonds to your fellow countrymen. We argue that the independent choice of a religion and its induced values are more important for a common identity than coincidental ethnic differences. The paper is organized as follows. Section 2 describes the calculation of our index of national identity. Data and methodology are presented in Section 3. Section 4 shows the empirical results, followed

by a discussion in Section 5. Section 6 briefly concludes.

2 National Identity Index

This section describes the construction of the index of national identity. The motivation behind constructing an index of national identity is twofold. First, the reason for constructing an index rather than analyzing several indicators related to national identity separately is that we aim to identify a common underlying factor captured by a set of indicators of national interests and orientations, namely the national identity of a person. Second, within a uni-dimensional index, we are directly able to analyze the determinants of national identity.

To derive the national identity index, we apply a principal component analysis. Principal component analysis is an aggregation technique to identify from a set of variables those linear combinations that best capture the common information behind the variables (Filmer and Scott, 2008). The main idea of this approach is to construct an aggregated uni-dimensional index over the range of different dichotomous indicators of national interests and orientations capturing the national identity of a person.

The approach of aggregating different variables to a uni-dimensional index is widely used in the economic and social literature. We closely follow the approach of Filmer and Pritchett (2001) and Sahn and Stifel (2001, 2003) to construct an index of material welfare based on the possession of housing durables. The authors propose an asset index based on the possession of household assets and dwelling characteristics as a proxy of material welfare of households in cases where no information on household income or expenditure are at hand. Paldam and Gundlach (2012) use an index approach do derive a measure of religiosity to analyze the religious transition over time.[2]

We assume that specific variables on attitudes towards the nation a person live in can explain the long-term national identity of a person measured by an aggregated index:

$$NI_j = b_1 a_{j1} + b_2 a_{j2} + ... + b_k a_{jk} \tag{1}$$

where NI_j is the national identity index, the a_j's refer to the respective variable of the person j recorded as dichotomous variables in the data and the b's are the respective weights for each variable used to aggregate the indicators to a one-dimensional index and that are to be estimated. In our model this

[2] A large body of literature also exists using an asset index to explain inequalities in educational outcomes (e.g. Ainsworth and Filmer 2006), health outcomes (e.g. Bollen et al. 2002), child mortality (e.g. Sastry 2004) when data on income or expenditure is missing. In addition, asset indexes are used to analyze changes and determinants of poverty (e.g. Stifel and Christiaensen 2007).

means that the k^{th} identity variable, identified by a_{jk} is a linear function of a common factor, which in our case is "national identity". We rely on the first principal component as our national identity index.[3]

We use micro data from the World Value Survey to calculate the index. The choice of the variables that enter the index depends on two factors: First, we use variables that are related to the national identity capturing available information in the World Value Survey on national interest and orientations. The second factor is mainly driven by data constraints. This means that we rely on a specific set of variables because it is identified as the set which suffers least from missing information. We include three dichotomous variables as components for the national identity index: the willingness of a person to fight for the country (1=yes), if the person is very proud of his/her country[4], and whether the person has the attitude that the country is the prime geographical group she belongs to (1=yes).

An issue that arises is the choice of the number of variables that enter the index. Including only three variables could be problematic in the sense that these variables do not capture the 'true' national identity, and that more variables are needed to capture the common factor. However, the first component explains around 30% of the covariance indicating that we capture the common factor behind the indicators fairly well. We also tried to derive the index based on more variables. The results of the index are quite similar, but since the sample size is then heavily reduced and since the results differ not very much, we decide to derive the index for as many countries as possible. With the underlying indicators, we are able to calculate the index for 69 countries in the sample, for over 170,000 persons.

Table 1: Descriptive statistics

Indicators	Mean	SD	Obs
Very proud of nationality (=1)	0.562	0.496	332,747
Willing to fight for the country (=1)	0.732	0.443	256,999
Geographical groups belonging to first: country (=1)	0.337	0.472	254,120
National identity index (mean)	-0.009		172,753
National identity index (sd)	1.009		
National identity index (min)	-1.921		
National identity index (max)	1.443		
% of the covariance explained by the first principal component	0.291		
Eigenvalue of first principal component	1.681		

Source: WVS; calculations by the authors.
Note: Indicators of national identity and index statistics.

Descriptive statistics on the indicators and on the index at the micro level are presented in Table 1. In the total sample, 56% of the respondents have stated that they are very proud of their country, 73% reported that they are willing to fight for their country, and 33.7% of the persons indicate that the country is the geographical group they feel they belong to. The mean value of the identity index is close to zero

[3] An alternative way to estimate the weights to derive the aggregated index is a factor analysis employed, for example, by Sahn and Stifel (2001) and Paldam and Gundlach (2012). However, the two estimation methods show very similar results. For a systematic overview of different aggregation techniques, see Filmer and Scott (2008).

[4] Based on the possible answers: very proud, quite proud, not very proud, not proud at all

with a range of around -2 to +1.5. Table A1 in the appendix shows the results by country for those countries where information on all variables that enter the index are available. The distribution of the index is presented in Figure A1 in the appendix.

3 Data and Methodology

3.1 Data

In the previous section we explained the methodology for constructing an index of national identity and described the indicators on which we rely. The principal component which we use to measure national identity will be the dependent variable throughout the whole analysis.

The explanatory variable of main interest is religious diversity, which is also called fragmentation or fractionalization. The index of religious diversity measures the probability that two randomly drawn people from a population belong to the same religious group. It is calculated as $1 - H$, where H is a Herfindahl-Index which is gained by $\sum_i^N s_i^2$, where s_i is there share of people belonging to each religious group i and N is the number of groups. We use the data on religious fractionalization provided by Alesina et al. (2003). We include ethnic diversity to control if our results are in line with the findings from Masella (2011). Information is taken from the same paper (Alesina et al., 2003).

We include income per capita to control for the possibility that economic development might have an impact on the formation of a national identity. Since income per capita might serve as a control variable for general economic development it can help reduce the negative consequences of possibly omitted variables. Hence income per capita is included in all regressions. In order to reduce the risk of reverse causality we want to use income per capita from a year before the observation period on national identity begins. The Maddison (2010) online database offers information on income per capita, also for the single former Soviet nations, for the year 1973.

The literature proposes further variables which might be correlated with national identity. Apparently, the political environment could influence identification with the nation. We control for the level of democracy by including the Polity score from the Polity IV database. It ranges from -10 for autocracies to +10 for full democracies. Freedom House offers slightly different measures. We use these indexes of political rights and civil liberties to take a closer look on democratic institutions. We rescale these two variables so that 1 stands for low political rights and civil liberties while the highest value is 7. We also generate a common factor of these two variables which can be understood as a measure of political participation (compare Gundlach and Opfinger, 2012).

Smith (1991) proposes that mobility throughout the country might also foster the formation of a

national identity. We take this possibility into account by controlling for physical, as well as, non-physical mobility. As a proxy for physical mobility we use the amount of kilometers of paved roads per 1,000 inhabitants. The data come from the CIA World Factbook. We use the number of phone lines per 100 inhabitants to capture the possible effect of non-physical mobility. Information is taken from the World Bank's World Development Indicators.

As it might be reasonable to assume that country size might affect national identity we include area size in our regressions. It is measured in square kilometers and the information is taken from the CIA World Factbook. Since we have a number of Latin American and African countries in our data sample we include a dummy variable which takes on the value 1 if a country has been under colonial rule. We rely, once more, on the country information in the CIA World Factbook.

Table 2 gives an overview of the summary statistics for the variables we use in our empirical estimations. Our measure of national identity is distributed between -1 and +1 on the country level. The highest value for national identity is calculated for Morocco whereas the lowest value results for Japan. Ireland comes closest to the mean value of 0.009, Bulgaria is the median observation.

Table 2: Country-level Summary Statistics

	Number of Observations	Mean	Median	Standard Deviation	Min	Max
National Identity	69	0.009	0.037	0.371	-0.939	0.886
Religious Diversity	69	0.447	0.458	0.218	0.004	0.860
Log Income 1973	69	8.526	8.58	0.871	6.210	9.810
Area in sqkm	69	1,249,469	238,391	3,051,825	316	17,098,242
Ethnic Diversity	69	0.340	0.320	0.226	0	0.930
Former Colony	69	0.304	0	0.464	0	1
Roads per 1,000	69	6.790	4.841	6.336	0.147	26.069
Phone Lines per 100	69	25.238	20.698	18.897	0.216	65.376
Polity sc. 1973	66	-0.894	-7	8.025	-9	10
Pol. Rights 1973	69	3.870	3	2.370	1	7
Civil Lib. 1973	69	3.900	3	2.276	1	7
Pol. Participation 1973	69	0.000	-0.385	1	-1.257	1.357

Morocco is the religiously least fractionalized country. The index of religious diversity reaches only 0.004 which means that the probability that two randomly drawn persons belong to different religions is less than half a percent. South Africa is the most diverse country with a probability of 86 %. With regards to ethnicity the most homogeneous country is South Korea. The index is virtually zero which means that more or less the whole population belongs to the same ethnic group. Uganda is the most fractionalized country in our sample with an index value of 0.93.

Income in 1973 is measured in 1990 international Geary-Khamis Dollars. Switzerland's income in 1973

was 9.81 logarithmic points which corresponds to 18,215 1990-Dollars. Income in Bangladesh amounts only to 6.21 logarithmic points which equals 498 1990-Dollars. The smallest country of our sample is Malta with a surface area of only 316 square kilometers. The largest country, Russia, covers an area of more than 17 million square kilometers.

Concerning physical mobility the infrastructure is worst in Tanzania where there are only 0.147 kilometers of paved roads per 1,000 inhabitants. The highest value is reported for Estonia. Uganda reveals the lowest value for phone lines per 100 inhabitants (0.22) which we use as a proxy for non-physical mobility. In Sweden there are more than 65 phone lines per 100 inhabitants.

In our sample there are 21 observations which are classified as former colonies. With regards to democracy, 17 countries out of our sample are classified as full democracies in 1973 in the Polity IV database. These are Western European countries and the Western off-shoots. No country is given a -10 for full autocracies, but five countries are ranked -9. A similar pattern emerges for the Freedom House variables although there are more countries in the categories representing the lowest political rights and civil liberties, respectively.

3.2 Methodology

The main goal of this study is to analyze the relationship between religious diversity and national identity. However, Alesina et al. (2003) report religious fractionalization for only one point in time for all countries and it probably does not change much in the time horizon from 1982 to 2005 for which we have data on national identity. Consequently, we are not able to apply panel data methods. Thus, we average our national identity index over the five WVS waves and run cross-country regressions. We use data at the country level because we want to analyze the impact of religious heterogeneity which is a societal phenomenon. We want to compare different societies and measure the impact of social heterogeneity on identity formation at the national level. In order to reduce the risk of reverse causality, most explanatory variables are used from 1973, the year for which we have income data for all countries.

The estimated model is of the form:

$$NI_i = \alpha + \beta \cdot reldiv_i + \gamma \cdot y_i + \delta X_i + \varepsilon_i$$

where NI_i is the index of national identity in country i, $reldiv_i$ is religious diversity in country i, y_i is income in country i, X_i is a vector of the other control variables and ε_i is the error term.

The coefficient of main interest is β. A positive coefficient would imply that national identity is higher in countries where the population is very diverse with regards to religion. A negative β would obviously mean that national identity is high if religious diversity is low. As presented in the Introduction we suspect

that higher levels of diversity should lead to lower levels of national identity. This means that we expect to find a negative relationship in our empirical estimations.

We have to take into account that religious fractionalization might suffer from endogeneity. Fincher and Thornhill (2008) propose that the disease environment can explain the differences in religious diversity. They argue that groups that share the same immunity pattern, i.e. they are immune to the same diseases, separate themselves from groups that have a different immunity pattern. This separation leads to a formation of different cultures, where religion is a part of. The boundaries which arise between the different groups prevents the flow of values and ideas so that the cultures between the groups evolve differently. Furthermore, Fincher and Thornhill (2008) can explain why religious diversity is higher in tropical regions than in temperate climatic zones. They explain that infectious diseases are more stressful in tropic environments. Consequently, "individuals disperse over shorter distances than where infectious disease is less stressful" (p.2). The result is that there are more distinct social groups in tropic areas which form different cultures and are grouped in one of nowadays' countries.

For the disease environment to be a suitable instrument the exclusion restriction requires that the disease environment does not affect the formation of a national identity through another channel than diversity. It is not reasonable to assume that people base their attitudes towards their countries on the probability of being affected by specific diseases. The variables that Fincher and Thornhill (2008) present describe the disease richness and the prevalence of pathogens. People who live in a country will most probably not base their national identity on these variables. The willingness to fight, pride for the nationality, and the preferred geographical group should be independent of the country's disease environment. Consequently, the exclusion restriction should not be violated so that we can use the variables on the disease environment as an instrument for religious diversity.

4 Results

This section presents the results of our empirical estimations. We regress our index of national identity on religious and ethnic diversity, the log of income per capita, country size, and a dummy for colonial past. We also include two different variables for mobility and four variables to capture the relationship between democratic institutions and national identity. Table 3 shows the results of the OLS regressions. Columns 1 and 2 use the Polity score as a measure of democracy, with physical mobility in column 1 and non-physical in column 2. The Polity score is substituted for political rights in columns 3 and 4, civil liberties in 5 and 6, and political participation in 7 and 8. These variables are combined once with physical and once with non-physical mobility. The t-statistics are shown in parentheses beneath the coefficients. *,**,*** denote

statistical significance at the 10, 5, and 1 percent level.

Table 3: OLS Results, dependent variable: National Identity

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Religious Diversity	-0.39	-0.34	-0.36	-0.31	-0.36	-0.33	-0.36	-0.32
	(-1.94)*	(-1.66)	(-1.85)*	(-1.58)	(-1.88)*	(-1.67)*	(-1.87)*	(-1.62)
Log Income 1973	-0.19	-0.14	-0.18	-0.14	-0.17	-0.14	-0.17	-0.14
	(-2.88)***	(-1.70)*	(-2.57)**	(-1.82)*	(-2.29)**	(-1.78)*	(-2.42)**	(-1.81)*
Area	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	(0.75)	(0.77)	(0.79)	(0.74)	(0.82)	(0.78)	(0.81)	(0.76)
Ethnic Diversity	0.12	0.06	0.07	0.01	0.06	0.01	0.06	0.01
	(0.61)	(0.29)	(0.34)	(0.05)	(0.32)	(0.03)	(0.32)	(0.03)
Former Colony	0.13	0.09	0.15	0.09	0.16	0.11	0.16	0.10
	(1.04)	(0.79)	(1.19)	(0.75)	(1.28)	(0.85)	(1.23)	(0.77)
Roads per 1,000	0.00		0.00		0.00		0.00	
	(-0.03)		(0.12)		(0.17)		(0.15)	
Phone Lines per 100		-0.01		-0.01		0.00		0.00
		(-1.16)		(-1.11)		(-0.87)		(-0.99)
Polity sc. 1973	0.00	0.01						
	(0.33)	(0.89)						
Pol. Rights 1973			0.00	0.02				
			(-0.18)	(0.62)				
Civil Lib. 1973					-0.01	0.01		
					(-0.44)	(0.24)		
Pol. Participation 1973							-0.02	0.03
							(-0.31)	(0.44)
cons	1.75	1.41	1.63	1.36	1.55	1.37	1.57	1.42
	(2.97)***	(2.15)**	(2.84)***	(2.21)**	(2.61)**	(2.19)**	(2.49)**	(2.22)**
Observations	66	66	69	69	69	69	69	69
R sq.	0.37	0.39	0.37	0.38	0.37	0.38	0.37	0.38

The explanatory variable of main interest is religious diversity. The sign of the coefficient is negative as we expected which implies that higher levels of religious diversity within a country are correlated with lower levels of national identity. However, it reaches statistical significance at the 10 percent level in only five cases, falling short of it in the other three regressions only slightly. The size of the coefficient lies between -0.3 and -0.4 which means that higher religious diversity by 0.1 index points relates to lower national identity by only 0.03 to 0.04 index points, a rather small effect.

The difference in religious diversity between the most homogeneous and most fractionalized countries amounts to 0.856 index points. This would imply a difference in national identity of 0.26 to 0.34 index points. An increase in religious diversity by one standard deviation corresponds to a lower level of national identity by 0.07 to 0.09 index points which is less than a quarter of a standard deviation.

Our results seem to support Masella's (2011) finding of no significant relationship between ethnic diversity and national identity. The index of ethnic diversity does not even come close to conventional levels of statistical significance. The coefficients are even slightly positive which runs counter the intuition that higher diversity is related to lower levels of national identity. But as already mentioned, the coefficients are very small and far from statistical significance.

Interestingly we find that higher levels of income are correlated with lower levels of national identity. This relationship is statistically significant at the 1 percent level in one regression, at the 5 percent level in three regressions, and at the 10 percent level in the remaining four. The size of the coefficient varies between -0.14 and -0.19. An income which is higher by one standard deviation corresponds to lower levels of national identity by between 0.12 and 0.17 index points. Comparison of the richest to the poorest country of our sample implies a difference of 0.5 to 0.68 index points in national identity.

Neither mobility nor the level of democracy seem to be related to our index of national identity. None of these variables comes close to statistical significance. The same applies to country size and the dummy variable for colonial past. The lack of statistical significance might be due to the sample size of only 66 observations in columns 1 and 2 and 69 observations in the remaining regressions. However the coefficients on mobility and democracy are also numerically very small and even switch signs in some cases. Apparently these variables are not significantly related to national identity. The coefficient on colonial past is similar in size throughout the different regressions. It might hint at a positive relationship between a colonial past of a country and national identity.

Table 4: IV Results, dependent variable: National Identity, Religious Diversity instrumented by disease and pathogen variables from Fincher and Thornhill (2008)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Religious Diversity	-1.33	-1.38	-1.33	-1.33	-1.37	-1.40	-1.35	-1.37
	(-2.04)**	(-1.84)*	(-2.23)**	(-2.10)**	(-2.27)**	(-2.13)**	(-2.25)**	(-2.10)**
Log Income 1973	-0.20	-0.20	-0.17	-0.18	-0.15	-0.17	-0.16	-0.17
	(-2.65)***	(-2.00)**	(-2.18)**	(-1.98)**	(-1.85)*	(-1.84)*	(-1.99)**	(-1.91)*
Area	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	(1.37)	(1.37)	(1.52)	(1.49)	(1.55)	(1.53)	(1.54)	(1.51)
Ethnic Diversity	0.41	0.43	0.33	0.33	0.33	0.36	0.33	0.35
	(1.42)	(1.26)	(1.26)	(1.15)	(1.28)	(1.19)	(1.26)	(1.16)
Former Colony	0.00	0.01	0.04	0.05	0.06	0.09	0.05	0.07
	(0.00)	(0.09)	(0.25)	(0.37)	(0.38)	(0.58)	(0.32)	(0.49)
Roads per 1,000	0.00		0.00		0.00		0.00	
	(-0.41)		(-0.20)		(-0.17)		(-0.18)	
Phone Lines per 100		0.00		0.00		0.00		0.00
		(-0.01)		(0.05)		(0.26)		(0.18)
Polity sc. 1973	0.00	0.00						
	(0.41)	(0.27)						
Pol. Rights 1973			-0.01	-0.02				
			(-0.59)	(-0.46)				
Civil Lib. 1973					-0.02	-0.03		
					(-0.86)	(-0.78)		
Pol. participation 1973							-0.05	-0.06
							(-0.74)	(-0.64)
cons	2.15	2.21	1.98	2.01	1.87	1.96	1.85	1.89
	(3.08)***	(2.41)**	(2.95)***	(2.54)**	(2.69)***	(2.47)**	(2.54)**	(2.43)**
Observations	66	66	69	69	69	69	69	69
R sq.	0.15	0.12	0.11	0.11	0.09	0.07	0.10	0.09

As we already discussed before, religious diversity might be subject to endogeneity issues arising from omitted variables or reverse causation with national identity. In order to tackle this problem we run

instrumental variable regressions. We instrument religious diversity with the disease and pathogen variables proposed by Fincher and Thornhill (2008). Both of these variables enter statistically significantly, in most cases at the five percent level, in the first stage regressions. The results of the second stages of the 2SLS regressions are presented in Table 4.

The instrumental variable regressions show an important difference to the OLS results. The coefficient on religious diversity is statistically significant in all eight regressions, in seven cases at the 5 percent level and at the 10 percent level in column 2. The relationship remains negative as we expected which supports the idea that higher levels of diversity correspond to lower levels of national identity. Apparently, the OLS regressions indeed suffer from endogeneity issues.

The magnitude of the coefficient is also significantly larger in absolute terms in the instrumental variable regressions compared to the OLS results. The coefficient ranges from -1.33 to -1.40 in column 6. This implies that national identity is by 0.133 to 0.14 index points lower in a country in which religious diversity is higher by 0.1 index points. The difference in religious diversity between the most homogeneous and most diverse countries amounts to a difference in national identity of 1.14 to 1.2 index points. Religious diversity that is higher by one standard deviation corresponds to lower national identity by 0.29 to 0.31 index points, or about 80 percent of a standard deviation.

Our instrumental variable results also support Masella (2011). We do not find a statistically significant relationship between ethnic diversity and the level of national identity. Again, the coefficient is positive which is counter-intuitive, but in any case it is not possible to reject the hypothesis of no relationship between these two variables.

Confirming the insights from our OLS estimations above, income seems to be negatively related to the level of national identity. The negative coefficient is statistically significant at the 1 percent level in column 1, at the 5 percent level in four regressions, and at the 10 percent level in the remaining three. The magnitude varies between -0.15 and -0.2 which is very similar to the OLS regressions.

The variables measuring mobility throughout the country and democracy remain far from statistical significance. Obviously there is no relationship between these variables and national identity. Also, the results on the dummy variable on colonial past is far less conclusive in the instrumental variable regressions. The size of the coefficients is only half of the coefficients from the OLS regressions. In column 1 of Table 4 the coefficient even switches to negative. It seems that there is no statistically significant relationship between a colonial past and national identity today.

Country size comes closer to statistical significance in the instrumental variable regressions. It falls short of significance at the 10 percent level rather slightly in all eight columns. However, this might again

be due to the rather small sample size. The coefficient appears to be zero. Dividing the surface area by one million to get interpretable coefficients reveals that a surface area that is larger by one million square kilometers is related to a level of national identity which is higher by about 0.03 index points which is an admittedly small effect.

In a nutshell, our results indicate that there is a negative relationship between religious diversity and the level of national identity. We can confirm the finding that ethnic diversity and national identity are not significantly related. Furthermore, our results suggest that the relationship between income and the level of national identity is also negative.

5 Discussion

We have created an index to measure national identity based on three indicators revealing people's attitudes towards their nation. These indicators are independent of the political institutions in the respective countries. Hence our index measures if the respondents feel a connection to their country or if they prefer to identify with other networks.

The empirical results in the previous section indicate that national identity is negatively related to religious diversity. It is a widespread hypothesis that diversity should have a negative impact on national identification. However, Masella (2011) reports that there is no significant relationship between ethnic diversity and national identity. We can confirm this finding, as ethnic diversity does not enter significantly in our regressions. Instead, we find a significant and negative relationship between our measure of national identity and religious diversity in Table 4.

Interestingly, Georgiadis and Manning (2009) find that religion is not important for national identity in Britain. However, the authors investigate only the role of each specific religious denomination. Our approach is different in that we are not interested in the role each religion plays, that is e.g. do Muslims or Christians identify more with the nation. We are interested in the relationship between the level of religious fragmentation in a society and national identity.

Contemplating our results in their entirety proposes that diversity in the society is indeed related to the formation of a national identity. However, this relationship is apparently not caused by ethnic differences but by religious fragmentation. Opfinger (2011) shows that religious diversity is also negatively related to the level of religiosity within a country. It seems that religious diversity is key in understanding the identity formation process within societies.

People search for networks with which they can identify. They generate utility from being part of a group that shares the same values and norms. These values and norms are to a large extent determined

by cultural beliefs and religion is an integral part of a culture. In countries that are religiously very homogeneous people find it easy to share a set of common values with their neighbors. They engage in the same rituals and obey the same norms. When they realize that the whole population follows these same rules a form of belonging together arises which results in the formation of a common national identity.

Societies that are constituted of many different religious groups find it harder to share a set of common values. It may be that the rituals of one religious group are regarded as strange or maybe even wrong from another religious group. This can lead to conflicts between religious groups within one society. Consequently, the adherents to different denominations separate themselves from each other and they do not form a common identity on the national level. This interpretation is in line with the findings of Bisin et al. (2010) who propose that cultural distinctiveness is the is more prevalent. Apparently, people who cannot identify with others on a broader base prefer to form distinct and different identities.

The role of religious diversity appears to be of major importance for the formation of a common identity compared to ethnic diversity. No one can be made responsible for his ethnic heritage. However, a religious network can be chosen independent of race. If people of different ethnicities engage in the same religious rituals and beliefs they share a common culture. On the contrary, different religious groups may find it hard to find compromises on which rituals and norms are right and which should be abandoned. Choosing a religious group is more important for the formation of a common identity than the given ethnic background. This can explain why ethnic diversity does not reveal a significant relationship with national identity while religious diversity does.

The other important finding which is worth being discussed is the relationship between income and national identity. Our results propose that higher income is related to lower levels of national identity. Apparently the way of causation is not definitely determinable in this setting. It is possible that higher national identity causes a decrease of income. Concentration on issues determining national identity might prevent people from working for higher income. The most extreme case would be starting a war with another country due to nationalist feelings. This would most probably lead to a reduction of income.

Nevertheless, the other way of causation intuitively appears to be more sensible. Rising incomes could lead to lower levels of national identity. Paldam and Gundlach (2012) show that rising levels of income lead to a reduction in religiosity. Gundlach and Opfinger (2012) and Hirschle (2011) argue that it becomes more costly to engage in time-consuming religious behavior as the income level rises. People face higher opportunity costs when spending several hours in a temple or church if they can generate a high income in the same time compared to a situation where wages are very low. This leads to a decrease in religiosity when income levels rise.

A similar way of argumentation may hold for national identity. If people face the opportunity to gain high income they might spend less time figuring out if their neighbors share the same values and engage in the same rituals. It might be more important to them with whom they can work together in order to generate income. Furthermore the social sciences and psychological literature proposes that there is a trend of isolation in the richest societies. This trend might be termed "rising individualism" (compare Kahneman, 2011). As the world becomes ever more globalized and networked via internet and other modern media the need for social interaction on a personal basis might decline. This induces people to reduce their interest in their neighbors and the rest of their fellow countrymen. In a highly individualistic society people do not feel the need to search for networks with which they can identify. This reduces the ties to one's society and therefore leads to a lower level of national identity.

As we discussed in the Introduction there is considerable evidence that a common identity is important for the functioning of a society. Consequently, the question arises what policy advise can be derived from our findings if governments pursue the goal of building a national identity. Apparently, reducing income in order to increase national identity is not a policy that should be pursued. However, a government might want to take measures to counteract the trend of rising individualism. This might be achieved, for example, by providing a variety of cultural offers so that people have incentives to interact.

More importantly is the role of religious diversity. Our results suggest that lower levels of diversity are related to higher national identity. However, it should not be the solution to force the whole population to adopt the same religion. As Georgiadis and Manning (2009) show people feel less connected to their nation if they feel they are not well treated or if they are discriminated against. Thus, governments should pursue policies that makes differences in religious affiliations less important in everyday life. The main policy goal should be to prevent people from being discriminated against because of their religion. Especially, if government policies are highly influenced by a religious group there arises the risk that certain minority groups feel excluded or disadvantaged. Due to the stringent relationship between the concepts of religion and national identity, it follows that governments may want to strictly separate church and state and promote a secular society.

6 Conclusion

We develop a measure for national identity to make it numerically comparable across countries. Our index is based on the respondents' attitudes towards their country, namely the willingness to fight for the country, pride to be a citizen of the country, and the geographical group you feel you belong to. These indicators are independent of the political circumstances so that our measure captures the national identity of a person.

We are able to calculate our index for 69 countries.

We investigate the relationship between national identity and diversity within a society. Earlier studies did not find a relationship between ethnic diversity and the level of national identity. We confirm this finding, however, our results indicate that religious diversity is negatively related to national identity. Our instrumental variable regressions reveal that higher levels of religious diversity by one standard deviation correspond to lower levels of national identity by about 80 percent of a standard deviation.

Furthermore, we find that income is also negatively related to our index of national identity. Other variables that were proposed to influence national identity, such as mobility throughout the country or democratic institutions, do not enter significantly in our regressions.

Our interpretation of the results on diversity proposes that people search for a network with which they can identify. If people belong to the same religion they share a set of common values and norms and engage in the same rituals. This makes it easy to identify with others. If the whole population belongs to the same religion they form a common culture and can also identify on a national basis. However, if people belong to different religions they cannot find a set of common values. High religious diversity may lead to conflict between the different groups which prevents the population from forming a common national identity.

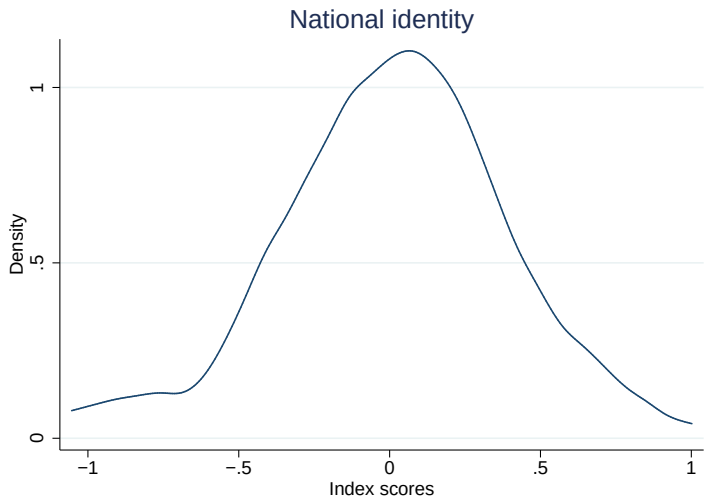
We propose that government policies should secure the rights of minorities so that people do not feel discriminated against as feeling excluded might weaken the ties to the nation. Furthermore it appears feasible that matters of the church should be separated from state and government issues.

Appendix

Table A1: National Identity on the country level (means)

Country	Index value	Willing to fight	Belonging to country	Proud of nationality	Country	index value	Willing to fight	Belonging to country	Proud of nationality
Albania	0.173	0.763	0.290	0.648	Lithuania	-0.349	0.750	0.279	0.272
Argentina	0.033	0.626	0.467	0.587	Luxembourg	-0.393	0.544	0.246	0.476
Armenia	0.063	0.804	0.455	0.435	Macedonia	0.309	0.815	0.369	0.643
Australia	0.320	0.710	0.438	0.706	Malta	0.235	0.721	0.306	0.759
Austria	-0.255	0.604	0.267	0.537	Mexcio	0.129	0.771	0.264	0.706
Azerbaijan	0.508	0.971	0.451	0.637	Modova	-0.212	0.774	0.332	0.257
Bangladesh	0.580	0.922	0.463	0.754	Morocco	0.886	0.883	0.617	0.776
Belarus	-0.157	0.880	0.270	0.312	Netherlands	-0.555	0.586	0.354	0.216
Belgium	-0.859	0.379	0.226	0.281	New Zealand	0.234	0.640	0.556	0.678
Bosnia and H.	-0.017	0.783	0.322	0.472	Nigeria	0.045	0.714	0.306	0.674
Brazil	-0.248	0.535	0.298	0.556	Norway	-0.056	0.893	0.176	0.479
Bulgaria	0.037	0.767	0.376	0.424	Peru	0.407	0.822	0.341	0.782
Canada	-0.029	0.640	0.303	0.656	Philippines	0.283	0.874	0.156	0.797
Chile	0.114	0.700	0.415	0.597	Poland	0.489	0.846	0.398	0.677
China	0.145	0.925	0.433	0.312	Portugal	-0.319	0.680	0.318	0.582
Croatia	-0.149	0.818	0.196	0.421	Romania	0.057	0.810	0.298	0.440
Czechia	-0.278	0.740	0.358	0.287	Russia	-0.317	0.807	0.230	0.333
Denmark	-0.152	0.809	0.265	0.395	Serbia	-0.151	0.756	0.233	0.396
Dominican Rep.	0.100	0.788	0.088	0.762	Slovakia	-0.388	0.706	0.210	0.320
Estonia	-0.364	0.832	0.171	0.251	Slovenia	0.292	0.863	0.375	0.580
Finland	0.080	0.852	0.372	0.481	South Africa	0.198	0.677	0.430	0.698
France	-0.470	0.577	0.286	0.344	Spain	-0.234	0.569	0.299	0.519
Georgia	0.241	0.714	0.493	0.696	Sweden	-0.059	0.884	0.245	0.403
Germany	-0.932	0.433	0.186	0.211	Switzerland	-0.360	0.712	0.212	0.345
Guatemala	0.377	0.696	0.405	0.878	Tanzania	0.467	0.931	0.249	0.817
Hungary	-0.052	0.799	0.304	0.553	Turkey	0.629	0.956	0.452	0.697
Iceland	0.159	0.755	0.493	0.601	Uganda	0.108	0.651	0.429	0.659
India	0.434	0.877	0.337	0.729	Ukraine	-0.316	0.778	0.263	0.257
Iraq	0.022	0.372	0.597	0.805	United Kingdom	-0.069	0.694	0.294	0.522
Ireland	0.021	0.611	0.282	0.725	Uruguay	0.158	0.548	0.520	0.747
Italy	-0.644	0.435	0.248	0.400	USA	0.236	0.737	0.291	0.745
Japan	-0.939	0.257	0.331	0.259	Venezuela	0.629	0.840	0.408	0.930
Korea (Rep. of)	-0.033	0.809	0.403	0.314	Vietnam	0.757	0.964	0.542	0.798
Krygyzstan	0.122	0.870	0.467	0.429	Zimbabwe	0.098	0.553	0.431	0.775
Latvia	-0.166	0.804	0.324	0.364					

Figure A1: Distribution of national identity on the microlevel



References

Ainsworth, Martha, Deon Filmer (2006). Children's Schooling: AIDS, Orphanhood, Poverty, and Gender. *World Development* 34(6), 1099-1128.

Akerlof, George A., Rachel E. Kranton (2000). Economics and Identity. *Quarterly Journal of Economics* 115(3), 715-753.

Alesina, Alberto, Arnaud Devleeschauwer, William Easterly, Sergio Kurlat, Romain Wacziarg (2003). Fractionalization. *Journal of Economic Growth* 8, 155-194.

Alesina, Alberto, Enrico Spolaore (2003). The Size of Nations. *Cambridge, MA: MIT Press.*

Anderson, Benedict (2006). Imagined Communities (2^{nd} ed.). *London: Verso.*

Austen-Smith, David, Roland G. Fryer, Jr. (2005). An Economic Analysis of 'Acting White'. *Quarterly Journal of Economics* 120 (2), 551-583.

Battu, Harminder, McDonald Mwale, Yves Zenou (2007). Oppositional identities and the labor market. *Journal of Population Economics* 20, 643-667.

Bisin, Alberto, Eleonora Patacchini, Thierry Verdier, Yves Zenou (2010). Bend It Like Beckham: Ethnic Identity and Integration. *NBER Working Paper, 16465.*

Bloom, William (1990). Personal Identity, National Identity and International Relations. *British Journal of Political Science* 20, 1-10.

Bodenhorn, Howard, Christopher S. Ruebeck (2003). The Economics of Identity and the Endogeneity of Race. *NBER Working Paper, 9962.*

Bollen, Kenneth A., Jennifer L. Glanville, Guy Stecklov (2002). Economic Status Proxies in Studies of Fertility In Developing Countries: Does the Measure matter?. *Population Studies* 56(1), 81-96.

Bond, Ross (2006). Belonging and Becoming: National Identity and Exclusion. *Sociology* 40 (4), 609-626.

Calhoun, Craig (1993). Nationalism and Ethnicity. *Annual Review of Sociology* 19, 211-239.

Checkel, Jeffrey T. (1999). Norms, Institutions, and National Identity in Contemporary Europe. *International Studies Quarterly* 43(1), 84-114.

Collier, Paul, Anke Hoeffler (1998). On economic causes of civil war. *Oxford Economic Papers* 50, 563-573.

Constant, Amelie F., Klaus F. Zimmermann (2008). Measuring Ethnic Identity and its Impact on Economic Behavior. *Journal of the European Economic Association* 6(2-3), 424-433.

Constant, Amelie F., Liliya Gataulina, Klaus F. Zimmermann (2009). Ethnosizing immigrants. *Journal of Economic Behavior and Organization* 69, 274-287.

Darity, William A.,Jr., Patrick L. Mason, James B Stewart (2006). The economics of identity: The origin

and persistence of racial identity norms. *Journal of Economic Behavior and Organization* 60, 283-305.

Esteban, Joan-María, Debraj Ray (1994). On the Measurement of Polarization. *Econometrica* 62(4), 819-851.

Fearon, James D., David D. Laitin (2003). Ethnicity, Insurgency, and Civil War. *American Political Science Review* 97(1), 75-90.

Filmer, Deon, Lant H. Pritchett (2001). Estimating Wealth Effects without Expenditure Data - or Tears: An Application to Educational Enrollments in States of India. *Demography* 38(1), 115-132.

Filmer, Deon, Kinnon Scott (2008). Assessing Asset Indices. *World Bank Policy Research Working Paper No. 4605.*

Fincher, Corey L., Randy Thornhill (2008). Assortative sociality, limited dispersal, infectious disease and the genesis of the global pattern of religious diversity. *Proceedings of the Royal Society B: Biological Sciences.*

Gillis, John R. (Ed.) (1996). Commemorations: the politics of national identity. *Princeton: Princeton University Press.*

Georgiadis, Andreas, Alan Manning (2009). One Nation Under a Groove? Identity and Multiculturalism in Britain. *CEP Discussion Paper No 944.*

Gundlach, Erich, Matthias Opfinger (2012). Religiosity as a Determinant of Happiness. *Review of Development Economics*, forthcoming.

Hirschle, Jochen (2011). The affluent society and its religious consequences: an empirical investigation of 20 European countries. *Socio-Economic Review* 9 (2), 261-285.

Jones, F.L. (1997). Ethnic diversity and national identity. *Journal of Sociology* 33(3), 285-305

Kahneman, Daniel (2011). Thinking, Fast and Slow. *New York City, NY: Macmillan.*

Maier, Charles S. (1997). The unmasterable past: history, holocaust and German national identity. *Cambridge, MA: Harvard University Press.*

Masella, Paolo (2011). National Identity and Ethnic Diversity. *Journal of Population Economics*, forthcoming.

Miles, William F.S., David A. Rochefort (1991). Nationalism Versus Ethnic Identity in Sub-Saharan Africa. *American Political Science Review* 85(2), 393-403.

Miller, David (2000). Citizenship and national identity. *Malden, MA: Polity Press.*

Montalvo, Jose G., Marta Reynal-Querol (2005). Ethnic diversity and economic development. *Journal of Development Economics* 76, 293-323.

Montalvo, Jose G., Marta Reynal-Querol (2005a). Ethnic Polarization, Potential Conflict, and Civil

Wars. *American Economic Review* 95(3), 796-816.

Noiriel, Gérard (1996). The French melting pot: Immigration, citizenship, and national identity. *Minneapolis: University of Minnesota Press.*

Opfinger, Matthias (2011). Religious Market Theory vs. Secularization: The Role of Religious Diversity Revisited. *University of Hannover Discussion Paper, 475.*

Paldam, Martin, Erich Gundlach (2012). The Religious Transition: A long-run perspective. *Public Choice*, forthcoming.

Sahn, David E., David Stifel (2001). Poverty Comparisons over Time and Across Countries in Africa. *World Development* 28(12), 2123-2155.

Sahn, David E., David Stifel (2003). Exploring Alternative Measures of Welfare in the Absence of Expenditure Data. *Review of Income and Wealth* 49(4), 463-489.

Sastry, Narayan (2004). Trends in Socioeconomic Inequalities in Mortality in Developing Countries: The case of Child Survival in Sao Paulo, Brazil. *Demography* 41(3), 443-464.

Smith, Anthony D. (1991). National Identity. *University of Nevada Press, Reno.*

Smith, Anthony D. (1992). National identity and the idea of European unity. *International Affairs* 68(1), 55-76.

Stifel, David, Luc Christiaensen (2007). Tracking Poverty Over Time in the Absence of Comparable Consumption Data. *The World Bank Economic Review* 21(2), 317-341.

Tajfel, Henri (1970). Experiments in intergroup discrimination. *Scientific American*, 223, 96-102.

Triandafyllidou, Anna (1998). National identity and the 'other'. *Ethnic and Racial Studies* 21(4), 593-612.

Vanhanen, Tatu (1999). Domestic Ethnic Conflict and Ethnic Nepotism: A Comparative Analysis. *Journal of Peace Research* 36(1), 55-73.

Wodak, Ruth, Rudolf de Cillia, Martin Reisigl, Karin Leibhart (1998). Zur diskursiven Konstruktion nationaler Identität. *Frankfurt/Main: Suhrkamp.*