IZA

DISCUSSION PAPER SERIES

IZA DP No. 4597

Dynamic Ethnic Fractionalization and Economic Growth in the Transition Economies from 1989 to 2007

Nauro F. Campos
Ahmad Saleh
Vitaliy Kuzeyev

November 2009

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Dynamic Ethnic Fractionalization and Economic Growth in the Transition Economies from 1989 to 2007

Nauro F. Campos
Brunel University, CEPR and IZA

Ahmad Saleh
Brunel University

Vitaliy Kuzeyev
Central European University

Discussion Paper No. 4597
November 2009

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 4597
November 2009

ABSTRACT

Dynamic Ethnic Fractionalization and Economic Growth in the Transition Economies from 1989 to 2007*

In their survey of the literature on ethnic fractionalization and economic performance, Alesina and La Ferrara (*JEL* 2005) identify two main directions for future research. One is to improve the measurement of diversity and the other to treat diversity as an endogenous variable. This paper tries to address these two issues: it investigates the effects of ethnic fractionalization on economic growth across countries using unique time-varying measures. We first replicate the finding of a weak effect of exogenous diversity on growth and then we show that accounting for how diversity changes over time and treating it as an endogenous variable makes a difference. Once diversity is instrumented (with lagged diversity and latitude), it shows a significant negative impact on economic growth which is robust to different specifications, polarization measures, econometric estimators, as well as to the use of an index of ethnic-religious-linguistic fractionalization.

JEL Classification: O11, Z12, O55, H1

Keywords: ethnic diversity, fractionalization, polarization, growth

Corresponding author:

Nauro F. Campos
Department of Economics and Finance
Brunel University (West London)
Uxbridge UB8 3PH
United Kingdom
E-mail: nauro.campos@brunel.ac.uk

* The authors would like to thank Jan Fidrmuc, Julius Horvath, Gabor Kezdi, Branko Milanovic, Jeffrey Nugent, James Robinson, Colin Rowat, seminar participants at Oxford University (CSAE) and two anonymous referees for valuable comments on previous versions, the Central European University for the hospitality and the Economic and Social Research Council for financial support (ESRC, Grant RES-000-22-0550.). The usual disclaimer applies. The ethnic, religious, and linguistic fractionalization (and polarization) indexes constructed for this paper are available on-line: http://www.naurocampos.net/delf19892007.xls

1. Introduction

There are three fundamental dimensions to any process of change. One is timing. When change starts and when it ends matters. Detecting the first signs that the status quo is sliding away is as difficult as identifying the moment when the previous status quo ceased to exist and the new one has fully established itself. The second fundamental dimension is extent. It refers to how much change was actually accomplished, whether the change itself was deliberate or unintended. The ratio between these first two dimensions is the speed of change. The third fundamental dimension is depth. This refers to how deep the effects of change turn out to be, whether or not the original change itself was intentional. There is no reason to think of these three dimensions as independent from each other. A case in point is that deep causes of change are often the most difficult ones to time, measure and attribute. Of course, this does not make them less important (the opposite is true, if anything). Institutions are a good example. They change slowly, over long periods of time, but their effects are widespread, long-lasting, and deep. We argue that the degree of fractionalization of a society along ethnic, religious or linguistic lines is in the same category. Fractionalization changes very slowly but this does not mean it does not change. It is also very difficult to measure but this does not mean they are short-lived. And diversity is often an extraordinarily deep phenomenon, but this does not mean we can afford to ignore it.

It was only in the last decade or so that ethnic fractionalization entered mainstream economics. There is now a burgeoning theoretical literature (see, e.g., Esteban and Ray, 1994, 1999, and Nehring and Puppe, 2002) and a very active empirical agenda. Although the seminal papers of Mauro (1995) and Easterly and Levine (1997) offer econometric evidence showing that greater levels of ethno-

linguistic fractionalization hinder economic performance, there has been less success in sustaining the evidence for such negative, direct effect. Easterly (2001) argues that the effect of ethnic fractionalization is conditional: it slows down economic activity only in countries with "sufficiently bad" institutions. Bluedorn (2001) and Alesina, Devleeschauwer, Easterly, Kurlat, and Wacziarg (2003) show that the negative impact of diversity on growth is particularly strong in less democratic countries. Posner (2004) argues that the negative effect is supported only by a restricted polarization index: restricted in that it includes only "politically relevant" ethnic groups. Montalvo and Reynal-Querol (2005a) argue that the direct effect of fractionalization on economic performance is weak and suggest focusing on polarization instead. In summary, the initial negative first-order effect of ethnic diversity on economic growth has been challenged and the literature seems to have turned to identifying the main channels through which diversity may affect the economy (i.e., indirect effects).

This large body of econometric evidence has two main features in common: diversity is measured using secondary data and diversity is often treated as a non-time varying, exogenous variable.[1] Alesina and La Ferrara (2005) provide an authoritative and thorough review of this empirical evidence and identify two main directions for future research: one is the need to improve the measurement of diversity and the other is the desirability of modeling diversity as an endogenous variable. The objective of this paper is to try to address these two issues. In this paper, we put together a data set

[1] These two features are related as the secondary data used to measure diversity refers to the early 1960s. The huge popularity of the index constructed by Soviet researchers and published in the *Atlas Narodov Mira* (Bruk and Apenchenko, 1964) is due in large part to its inclusion in Taylor and Hudson's *World Handbook of Political and Social Indicators (1972)*. For studies that use this index, see Mauro (1995), Easterly and Levine (1997), Collier (2001), La Porta, Lopez-de-Silanes, Shleifer and Vishny (1999) and Woo (2003a, 2003b).

that contains mostly primary, census-based, data. In terms of treating diversity as an endogenous variable, we make use of the genuine time variation shown by these indexes that, to the best of our knowledge, is unique to our data set. We propose lagged diversity and latitude as the instrument set and subject these to a comprehensive series of diagnostics tests (which they pass).

In what follows we report on the construction of a unique data set based mostly on primary data (national censuses) to measure ethnic diversity over time for a sample of countries that closely resemble a "natural experiment" (the 26 former centrally-planned economies, from 1989 to 2007).[2] These are said to resemble a "natural experiment" because until 1989 they shared a very similar set of economic and political institutions (central planning under socialism), but have since followed radically different economic and political trajectories. Using these data, we are able to replicate the most recent results from the literature and show that static (exogenous) diversity is indeed not robustly correlated with economic growth. However, when we capture empirically how ethnic diversity changes over time and model it as an endogenous variable, we conclude that ethnic fractionalization is negatively related to growth and this is robust to the use of different econometric estimators, specifications, polarization measures as well as to an index of ethnic-linguistic-religious fractionalization.

The rest of the paper is organized as follows. Section 2 discusses the data collection effort and the measurement methods used. Section 3 discusses the econometric methodology, presents the main results and subjects them to various robustness tests. Section 4 concludes with some brief suggestions for future research.

[2] Campos and Kuzeyev (2007) examine the relationship between growth and diversity between 1989 and 2002 within an endogenous growth framework, while this paper uses the traditional Solow model to study the growth-diversity nexus between 1989 and 2007.

2. Measurement

In this paper, we collect primary data (census-based) to measure ethnic diversity (fractionalization and polarization) over time (from 1989 to 2007) for a sample of 26 former centrally-planned economies.[3] National censuses are the preferred and most reliable source of ethnic diversity data. Unfortunately they are only conducted once a decade, at best. Micro-censuses and demographic surveys, which are arguably the second best sources of primary diversity data, tend to be conducted at five-year intervals. With this in mind, we assess what would be the maximum number of time periods for which we could obtain a balanced panel data set on the demographic (ethnic) composition of these 26 transition countries. We identify four time periods: 1989-1993, 1994-1998, 1999-2002, and 2003-2007. This means we use primary data from national censuses for the first and third period and data from micro-censuses and demographic surveys for the second and fourth period.[4]

We collect data on the percentage of the population belonging to each ethnic group in each country for each of these four periods. This generates a panel with 104

[3] We divided the sample in five groups for exposition purposes (Figure 1). The transition countries in ASIA are Armenia, Azerbaijan, Georgia, Kazakhstan, Kyrgyztan, Mongolia, Tajikistan, Turkmenistan and Uzbekistan. The BALKAN countries are Albania, Bulgaria, Croatia, Macedonia, Moldova and Romania. The BALTIC countries are Estonia, Latvia and Lithuania. The group called BUR comprises Belarus, Ukraine and Russia. The VISEGRAD countries are the Czech Republic, Hungary, Poland, Slovakia and Slovenia. CEEB stands for Central and Eastern European and Baltic countries and which is the sum of the BALTIC, BALKAN and VISEGRAD sub-groups.

[4] Although it is difficult to objectively judge the quality of these different sources of data, note that in each country collection of these data was done by the same agency, with comparable methodologies. They differ in that censuses cover the entire population and micro-censuses cover a representative sample. These figures were checked against various additional sources, including Rosenko (1999) *Nasii I Etnosi V Sovremennom Mire* (*Nations and Ethnicity in Today's World*) and *Natsionalniy Sostav Naseleniya SSSR* (*Ethnic Composition in The USSR*,

observations. Census data are available for about half of the observations. Once all the data were collected, we note that for some countries there were more than one estimate for a given time period, so a decision rule was needed. If two or more sources gave identical information up to the third decimal place, we first single out these sources. From them, we chose the combination that gave the most balanced distance among the indices over time. This was done to have the largest possible time span within the sample periods.[5] If we still have a tie, that is, if the remaining sources diverged up to the second decimal place, we used the one that caused less variability of the indices for the country in question over time. This rule of most balanced distance attempts to minimize source-variability bias as much as possible.

For the computation of the fractionalization indices, we apply the commonly used formula capturing the probability that two randomly selected individuals belong to different ethnic groups (e.g., Alesina et al. 2003):

$$F = 1 - \sum_{i}^{n} s_i^2 \qquad (1)$$

where s_i is the share of total population belonging to ethnic group *i*. The index takes values between zero (for a perfectly ethnically homogeneous country) and one (highly heterogeneous country).

One shortcoming of this measure is that the same value of the index can correspond to different distributions (Fearon, 2003). This sensitivity of the index to the total number of underlying groups requires attention. We compare two approaches. First, we use an unrestricted set with all disaggregated data allowing the

---

Finansi I Statistika, 1991).

[5] For example, we found data on the ethnic composition of the population in Latvia for the years 1994, 1995 and 1996 from different sources, whose indices were identical up to third decimal place. Hence, according to our rule, the time series 1989-1994-2000 was preferred to 1989-1996-2000.

number of ethnic groups for each country to vary over time. In the second approach, we restrict the number of groups for each country to be the same over time.[6] We find the differences are small.[7]

Figure 1 shows that these countries end up much more ethnically homogenous than they started with over a short period of time. This suggests that there may be value in re-thinking the assumption of exogeneity. Why does diversity change over time? One general cause is, of course, migration flows. These may be driven by better economic performance and opportunities in the destination country as well as by inferior economic performance and/or civil war and ethnic cleansing in the origin country. In developing countries, such a process should surely take decades to unfold. However, there are special circumstances in our sample of transition countries which allow for this process to take place in a much shorter period of time. Firstly, with the collapse of communism, workers become free to move to other countries (while under communism mobility restrictions often referred to the city, let alone country) in search

---

[6] The average number of ethnic groups in the restricted sample was 5.19 and in the unrestricted sample 7.04. Alesina et al.'s and Fearon's analogous figures for Eastern Europe and former Soviet Union countries are 6.48 (27 observations) and 4.55 (31 observations), respectively. The lowest number of groups in our data, including "others", is 3 (in several cases), while the largest is 8 (12 for Mongolia in the unrestricted sample). In addition to data quality, we must also be concerned with data comparability. In this respect, there are few dimensions over which researchers can exert some control. One of the few, however, refers to the number of groups used in the computation of the diversity indexes. Here we explore different ways of using this information across countries and over time. We find that these variations do not affect our main conclusions.

[7] For instance, the mean of this ethnic fractionalization index declines from 0.3726 (0.3768) in the first period to 0.345 (0.3538) in the second period to 0.3147 (0.3154) in third period to 0.30145 (0.30314) in the fourth period (values using the unrestricted number of groups are in parenthesis). For comparison, Alesina et al.'s value for the early 1990s is 0.3696, while Fearon's is 0.3723.

of better economic opportunities (Campos and Coricelli, 2002). Secondly, the ubiquitous Russian minorities seem to have been made to feel unwelcome and the new economic and political situation after 1991 results in return migration, causing the share of Russians to fall in every country in our sample, with the exception of Moldova. It is only after 1945 that Russians become the second largest ethnic group in most of the Republics (one example is Kazakhstan, where the national census of 1989 shows that the shares in total population are 37.8% to 39.7% for Russians and for Kazakhs, respectively.) A third important factor is violent conflict, for example, the wars in the Caucasus and former Yugoslavia. Because of the latter, for example, the share of Serbs in Croatia declines from 12.2% in 1991 to 4.54% in 2001.

Another concern about the existing ethnic fractionalization indices is that the definition of ethnic groups may change for political reasons. Alesina et al. (2003) remark that Somalia was counted as a homogeneous country prior to the civil war in 1991 with the notion of linking clans to ethnic groups coming into being only after that. Note that there are no disputes about group definitions in our data. Census questionnaires enumerate a fixed number of ethnic groups and let the respondent indicate to which she belongs. The residual option of "others" or "none of the above" is provided and taken into account (as one single group) when computing our diversity measures.

The emerging consensus is that polarization is the theoretically appropriate concept for measuring diversity.[8] The family of polarization measures developed by Esteban and Ray (1994, 1999) has been implemented in various ways. The one proposed by Alesina et al. (2003) is as follows:

[8] Fractionalization measures increase in the number of groups, while polarization maximum is reached with two groups of equal size.

$$ADEKW \quad Polarization = K\sum_{i=1}^{n}\sum_{j=1}^{n} s_i^{1+\alpha} s_j \left| y_i - y_j \right| \tag{2}$$

where K is a scaling factor and α is a constant between 0 and 1.6. Note that this formulation requires a measure of distance between groups (the last term in the right-hand side). Conceptually, distance can be thought of, for instance, as differences in median incomes. Because of data constraints, distance is often assumed to be constant.[9]

An alternative, yet related, implementation is the one proposed by Montalvo and Reynal-Querol (2005):

$$MRQ \quad Polarization = 1 - \sum_{i=1}^{n} \left(\frac{0.5 - s_i}{0.5} \right)^2 s_i \tag{3}$$

Notice that although Esteban and Ray (1994) and Montalvo and Reynal-Querol (2005) may look similar they are rather different.[10] Esteban and Ray deal with the calculation of polarization when distances are continuous while Montalvo and Reynal-Querol provide an index to calculate polarization when distances are discrete.

We use equations (1), (2), and (3) to calculate various measures of fractionalization and polarization. Appendix 1 shows that the pair-wise correlations between our measures, on the one hand, and investment, human capital and labor growth rates, on the other, is small (the largest is .11). Notice that the simple correlation among our measures of fractionalization and polarization is high (the smallest is .83). It is also worth noting that while the correlation coefficients between our diversity measures and human capital tend to be positive, the same with respect to investment and population growth tend to be negative (although in both cases they are

[9] To be more precise, the ADEKW index of polarization is the original index of polarization of Esteban and Ray (1994). The Alesina et al. (2003) index are obtained using different values of α and under the assumption that distance is constant and equal to 1.

not statistically significant). In between these extremes, the negative correlation between growth and all our measures of fractionalization and polarization is milder, ranging from -.24 to -.37.

3. Results

The objective of this section is to revisit the effect of ethnic diversity on economic growth. To do so, we estimate the standard augmented Solow model proposed by Mankiw, Romer and Weil (1992).[11] MRW's econometric specification is as follows:

$$\ln[\frac{Y}{L}] = \beta_0 + \beta_1 \ln s_k + \beta_2 \ln s_h - \beta_3 \ln(n + g + \delta) + u \qquad (4)$$

where Y/L is output per capita, s_k is the rate of investment in physical capital, s_h is the rate of investment in human capital, n is the population growth rate, g is the rate of technological change and δ is the depreciation rate.[12] Subscripts for countries and (the four) time periods are omitted. Notice that although the estimation in the original MRW paper was done by OLS, we here follow the more recent literature (e.g., McCleary and Barro, 2006) and first estimate (4) using SUR.[13]

Table 1 has our results treating polarization and fractionalization in a manner similar to that of the literature, that is, as exogenous variables. The specifications in

---

[10] We thank an anonymous referee for this point.

[11] In the fractionalization and growth literature, this approach is used by, for instance, Montalvo and Reynal-Queyrol (2003).

[12] We follow MRW in assuming that the sum of g and δ is constant. Although they assume it is constant at 5%, here we report results assuming that the sum of rates is 7.5% so as to reflect the larger depreciation observed in the capital stocks inherited from the socialist period.

[13] In previous versions of this paper we also reported specifications for the level of per capita GDP in the left-had side, instead of the growth rate. Our main results are unaffected by this change. In other words, we still find that fractionalization is important vis-à-vis growth only when treated as an endogenous variable and its dynamics is taken into account.

Table 1 all include initial income while all of those in Table 2 exclude it. Column 1 in Table 1 shows that the coefficients on investment, human capital, and population carry their expected signs (positive, positive and negative, respectively). Initial income is negative, but insignificant.[14] Exogenous ethnic fractionalization, however, has an almost negligible effect on growth.[15] Column 2 shows that the ethnic fractionalization index has no effect on growth, while columns 3 and 4 show that diversity is also not significant when proxied by any of the two versions of the Alesina et al.'s polarization measure. The same conclusion holds for the Montalvo and Reynal-Querol measure (column 5). These results may well be driven, for example, by ethnicity not being the appropriate dimension for conflict in these countries. In order to address this possibility, we computed two additional indexes. First, a principal components index of ethnic, linguistic and religious fractionalization dimensions was constructed. Column 6 reports these results and shows that this broader index is also not statistically significant. Second, we constructed an average index of these three fractionalization dimensions. Column 7 confirms that diversity is still statistically insignificant.

Table 2 repeats the estimations of Table 1 but excluding initial income. Column 1 shows that the coefficients on investment, human capital, and population are now all significant and carry their expected signs (positive, positive, and negative, respectively). However, the coefficients of all diversity indexes are still not significant, except column 4 which shows that the ethnic polarization index with

[14] There are important data quality issues that should be kept in mind when interpreting these results (for a discussion see Campos and Coricelli, 2002).

[15] The results from a standard Granger-causality test show that there is no evidence supporting the notion that growth (Granger-) causes diversity. These are available from the authors upon request.

α=1.6 is negative and statistically significant.

In sum, these findings on diversity are in line with most of the recent literature in that these estimates show that its direct effect on economic performance is weak. One possibility that the literature has not yet explored is that diversity changes over time and may also be endogenous (see, e.g., Alesina and La Ferrara, 2005). Fast growing economies will attract migrants, while newly independent states may try to expel formerly dominant ethnic minorities (say, Russians). We now turn to econometric results that try to take these possibilities into account.

In Table 3, we report our estimates of the augmented Solow model using instrumental variables techniques. These allow us to treat ethnic diversity as an endogenous variable. With initial income in the specification, we find that the coefficients on investment, population, and initial income are not significant, although the one for human capital is positive and statistically significant. Column 1 shows our results using dynamic (endogenous) diversity: the coefficient on ethnic fractionalization is now negative and significant. Columns 2 and 3 show that for the two versions of the Alesina et al.'s polarization measure and for the Montalvo and Reynal-Querol's index (column 4), the coefficients on ethnic fractionalization are also negative and now statistically significant. Column 5 shows that our principal components' ethnic-linguistic-religious fractionalization index generates similar conclusions, namely that dynamic and endogenous fractionalization seem to have a negative and robust first-order effect on economic growth. Column 6 presents similar results for the average of the three diversity dimensions (ethnic, religious and linguistic).

Table 4 presents similar results but excluding initial income from all specifications. As it can be seen, the only standard explanatory variable that remains

significant is human capital. For the set of diversity indexes, we can see that all of the relevant coefficients are still negative and all except one (out of six) are statistically significant.

Instrument selection is always a difficult matter. It is made more severe in this case by the fact that there has been little effort to explain theoretically or empirically the evolution of ethnic diversity over time. In this light, we tried a number of variables and combinations of variables and settled on the lagged diversity index and latitude (the absolute value of distance from the equator). We subject this choice to extensive testing and conclude that these two variables perform satisfactorily.

We start by examining the Sargan-Hansen test of overidentifying restrictions. The objective of this test is to help establish the validity of the instruments, that is, that the instruments are uncorrelated with the residuals and that their selection is justified. A rejection of the null hypothesis would suggest that the instruments are not valid. As it can be seen in the "diagnostics" panels of Tables 3 and 4, the null is rejected in all cases at conventional (95%) confidence levels suggesting that these instruments are indeed valid.

In terms of identification, next we report on tests for the relevance of the instrument set, that is, whether the instruments are correlated with the endogenous regressors. We report the Shea Partial R-squared (with only one endogenous regressor, this statistic is equivalent to the more common partial R-square) and the F-test of the excluded instruments in the corresponding first-stage regression. The results from these two tests support the validity of our set of instruments. The R-square figures are very high and the value of the F-statistic is above 10 in all specifications of tables 3 and 4. The Anderson canonical correlation likelihood-ratio test (CCLR) corroborates these conclusions.

Finally, we also report the Pagan-Hall and RESET tests. The Pagan-Hall tests for heteroskedasticity in the IV context. Given the extraordinary variation in growth performance across these transition economies over time, some may worry that this can be an important source of bias. None of the results in tables 3 and 4 suggest heteroskedasticity problems in the estimated equations' disturbance process. The RESET test we report is the Ramsey's regression error specification test as proposed by Pesaran and Taylor (1999). It shows all models in tables 3 and 4 are correctly specified in that omitted variables bias does not seem to be severe.

For the sake of sensitivity analysis, we also apply the Blundell and Bond's (1998) System GMM estimator. Table 5 presents GMM estimations for our augmented Solow model. The coefficients on investment, population, and human capital are statistically insignificant in all specifications (Table 5). The diversity indices are treated as exogenous and it can be seen that none of the various versions of the index is significant (the fractionalization index in Column 1, the Alesina et al. polarization index in column 2, the MRQ index in Column 3, and in columns 4 and 5, our two ethnic-linguistic-religious fractionalization indices). Notice that we were not able to generate results for the Alesina et al. 1.6 polarization index. Overall, these results in table 5 are similar to the one we discussed above in that exogenous diversity has no discernible first-order impact on economic growth.

Table 6 reports System GMM results when diversity is treated as endogenous. In this case, the coefficients on human capital are positive and now statistically significant. Population and investment are insignificant. Once the diversity indexes are treated as endogenous, the coefficients all carry the hypothesized negative sign and are statistically significant. The instruments applied are the lagged dependent variable and the latitude variable. The system GMM estimator uses as instruments for

the original equation, the first difference of all variables, while for the differenced equation, instruments are the lagged variables of the original model. In our case, investment, human capital, and population are considered as predetermined explanatory variables which are expected to be not correlated with the past and present value of the errors, while latitude is considered strictly exogenous. Notice that the test for the first-order residual serial correlation suggests that the model does not suffer from serial correlation.[16] Moreover, the validity of the instruments in the system GMM results is supported by the Hansen test. Note also that we use the two-step estimation, where the standard errors are corrected for panel specific auto-correlation and heteroscedasticity and we also apply the Windmeijer correction (Roodman, 2006). In our view, we prefer the results in tables 3-4 to those in Tables 5-6 (that is, we prefer the IV estimates to the System GMM ones) because our panel is very short both in terms of countries and especially in terms of time periods. Despite the potential problems with the GMM results, it is comforting to see that the main conclusions change little compared to the IV results, namely, that exogenous and static diversity seem to have little effect on growth while the same effect is much more statistically robust and economically meaningful from a model that takes into account the dynamics of diversity as well as of its exogenous component.

4. Conclusions

This paper investigates a number of questions related to the behavior of ethnic diversity over time and across countries and its effects in terms of economic

[16] As our panel covers only four periods and we use the one-period lagged diversity as an instrument, we are unable us to run the AR(2) test.

performance. We studied how much weight should be attached to the assumption that ethnic diversity does not change over time. We noted that this assumption is used widely. Paradoxically, the index of ethnic fractionalization that is commonly used in the literature was developed by researchers from former communist countries, that turn out to experience most dramatic changes in ethnic diversity in a very short period of time. We use census or micro-census data to create such indices for four points in time for a sample of 26 transition economies. Using these measures, and in line with the recent literature, we find weak evidence of a direct effect of diversity in the standard augmented Solow growth model. On the other hand, our panel estimates show that dynamic (endogenous) ethnic fractionalization is negatively related to growth, with equally robust results obtaining for measures of ethnic polarization.

As mentioned above, there are a number of issues that make the situation of ethnic groups in the transition countries somewhat special. In our view, those reasons support the dramatic changes in the ethnic composition we observe in such a short period of time. Although we do not think it is reasonable to expect that changes of this magnitude could be observed for other groups of developing countries over ten years or so, data may be available that would allow future research to relax the assumption that since 1960, that is over the last half-century, the degree of ethnic homogeneity has not change meaningfully in poorer countries. Such test can be accomplished, for instance, using decade averages of available ethnic diversity measures. This will be useful in re-assessing the recent discussion about the channels through which diversity (indirectly) affect growth. It is clear, however, that the construction of census-based measures for larger samples of developing countries over longer periods of time is still a rather demanding task.

References

Alesina, A., A. Devleeschauwer, W. Easterly, S. Kurlat and R. Wacziarg, "Fractionalization," *Journal of Economic Growth*, 8, 155-194, 2003.

Alesina, A. and E. La Ferrara, "Ethnic Diversity and Economic Performance," *Journal of Economic Literature*, 43 (3), 762-800, 2005.

Bluedorn, J., "Can Democracy Help? Growth and Ethnic Divisions," *Economics Letters*, 70, 121-126, 2001.

Bruk, S. I. and V. S. Apenchenko, eds. *Atlas Narodov Mira*. Moscow: Glavnoe upravlenie geodezii i kartografii gosudarstvennogo geologicheskogo komiteta SSSR and Institut etnografii im. H. H. Miklukho-Maklaia, Akademiia nauk SSSR, 1964.

Campos, N. and F. Coricelli, "Growth in Transition: What We Know, What We Don't, and What We Should," *Journal of Economic Literature*, XL (3): 793-836, 2002.

Campos, N. and V. Kuzeyev, "On the Dynamics of Ethnic Fractionalization," *American Journal of Political Science* 51(3): 620-639, 2007.

Collier, P., "Ethnic Diversity: An Economic Analysis," *Economic Policy*, 32, 127–166, 2001.

Easterly, W., "Can Institutions Resolve Ethnic Conflict?," *Economic Development and Cultural Change*, 49 (4) 687-706, 2001.

Easterly, W. and R. Levine, "Africa's Growth Tragedy: Policies and Ethnic Divisions," *Quarterly Journal of Economics*. CXII (4), 1203-1250, 1997.

Esteban, J. and D. Ray, "On the Measurement of Polarization," *Econometrica*, 62: 4, 819–851, 1994.

Esteban, J. and D. Ray, "Conflict and Distribution," *Journal of Economic Theory*, 87, 379–415, 1999.

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

La Porta, R., F. Lopez-de-Silanes, A. Shleifer and R. Vishny, "The Quality of Government," *Journal of Law, Economics and Organization*, 1999.

Mankiw, G., D. Romer and D. Weil, "A Contribution to the Empirics of Economic Growth," *Quarterly Journal of Economics*, 107: 2, 407-437, 1992.

Mauro, P., "Corruption and Growth," *Quarterly Journal of Economics*, 110:3, pp. 681-712, 1995.

McCleary, Rachel and Robert Barro, "Religion and Economy," *Journal of Economic Perspectives* 20 (2): 49-72, 2006.

Montalvo, J. and M. Reynal-Querol, "Religious Polarization and Economic Development," *Economics Letters*, 80: 201–210, 2003.

Montalvo, J. and M. Reynal-Querol, "Ethnic Polarization, Potential Conflict and Civil Wars," *American Economic Review*, 95: 796-816, 2005.

Pesaran, M.H. and Taylor, L.W., "Diagnostics for IV Regressions," *Oxford Bulletin of Economics and Statistics*, Vol. 61, No. 2, pp. 255-81, 1999.

Roodman, D., "How to Do xtabond2: An Introduction to Difference and System GMM in Stata," Working Paper 103. Center for Global Development, Washington, 2006.

Rosenko, M. (ed.), *Nasii i etnosi v sovremennom mire (Nations and ethnicity in today's world)*, Reference-Dictionary, S-Petersburg, 1999.

Finansi I Statistika (Soviet Ministry of Science) *Natsionalniy sostav naseleniya SSSR (Ethnic composition in USSR)*, Moscow, 1991.

Nehring, K. and C. Puppe, "A Theory of Diversity," *Econometrica*, 70: 3, 1155-98, 2002.

Posner, D., "Measuring Ethnic Fractionalization in Africa," *American Journal of Political Science*, 48:4, 2004

Taylor, C. and M. Hudson, *World Handbook of Political and Social Indicators*, 1972.

Woo, J., "Social Polarization, Industrialization, and Fiscal Instability: Theory and Evidence," *Journal of Development Economics*, 72: 223-252, 2003a.

Woo, J., "Economic, Political, and Institutional Determinants of Public Deficits," *Journal of Public Economics*, 87: 387–426, 2003b.

Table 1
SUR Estimation of augmented Solow model with exogenous diversity indices
(Dependent variable: growth rate of per capita GDP)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Ln I/Y	0.0620	0.0923	0.0990	0.111	0.101	0.0848	0.0683
	(0.0943)	(0.0989)	(0.101)	(0.105)	(0.102)	(0.100)	(0.102)
Ln HK	0.128*	0.141**	0.142*	0.142*	0.143*	0.143**	0.112
	(0.0688)	(0.0713)	(0.0725)	(0.0744)	(0.0732)	(0.0727)	(0.0785)
Ln(n+ g+ δ)	-0.0109	-0.0108	-0.0147	-0.0213	-0.0164	-0.0145	-0.00896
	(0.0374)	(0.0383)	(0.0389)	(0.0402)	(0.0392)	(0.0387)	(0.0412)
Ln Initial income	-0.00111	-0.00299	-0.00176	-0.00141	-0.00144	-0.000477	0.00272
	(0.0443)	(0.0464)	(0.0471)	(0.0485)	(0.0477)	(0.0472)	(0.0486)
Ethnic Fractionalization		-0.150					
		(0.142)					
Ethnic Polarization (α=0.8)			-0.414				
			(0.379)				
Ethnic Polarization (α=1.6)				-0.934			
				(0.833)			
Polarization (MRQ)					-0.126		
					(0.118)		
Ethno-linguistic-religious fractionalization						-0.006	
						(0.0215)	
Ethno-linguistic-religious fractionalization (average)							-0.003
							(0.00217)
Constant	-0.347	-0.268	-0.267	-0.234	-0.268	-0.369	-0.258
	(0.497)	(0.525)	(0.535)	(0.560)	(0.542)	(0.527)	(0.556)
Observations	24;24	24;24	24;24	24;24	24;24	24;24	24;24
	24;24	24;24	24;24	24;24	24;24	24;24	24;24
R-squared	-0.39;0.84	-0.32;0.84	-0.33;0.84	-0.34;0.84	-0.34;0.84	-0.45;0.83	-0.31;0.83
	0.77;0.82	0.77;0.82	0.77;0.83	0.76;0.83	0.76;0.83	0.77;0.83	0.76;0.83

Note: SUR estimates. Standard errors in parentheses, * indicates significant at 10%, ** significant at 5%, *** significant at 1%.

Table 2
SUR Estimation of augmented Solow model with exogenous diversity indices and without initial income
(Dependent variable: growth rate GDP per capita)

	(1)	(2)	(3)	(4)	(5)	(6)
Ln I/Y	0.576***	0.588***	0.586***	0.614***	0.589***	0.539***
	(0.143)	(0.145)	(0.146)	(0.142)	(0.146)	(0.151)
ln HK	0.314***	0.304**	0.284**	0.252**	0.276**	0.294**
	(0.117)	(0.120)	(0.119)	(0.112)	(0.119)	(0.129)
ln(n+ g+ δ)	-0.130**	-0.121*	-0.131**	-0.145**	-0.132**	-0.131*
	(0.0628)	(0.0643)	(0.0634)	(0.0589)	(0.0630)	(0.0669)
Ethnic Fractionalization		-0.369				
		(0.290)				
Ethnic Polarization(α=0.8)			-0.949			
			(0.724)			
Ethnic Polarization(α=1.6)				-2.669*		
				(1.377)		
Polarization (MRQ)					-0.298	
					(0.217)	
Ethno-linguistic-religious fractionalization						-0.0552
						(0.0391)
Constant	-0.358	-0.152	-0.0972	0.146	-0.0562	-0.358
	(0.535)	(0.558)	(0.560)	(0.544)	(0.560)	(0.571)
Observations	24;24;24;24	24;24;24;24	24;24;24;24	24;24;24;24	24;24;24;24	24;24;24;24
R-squared	-3.58;0.19	-3.13;0.22	-2.98;0.23	-2.87;0.25	-2.94;0.24	-2.85;0.22
	0.16;0.32	0.17;0.34	0.17;0.34	0.17;0.38	0.17;0.34	0.17;0.33

Note: SUR estimates. Standard errors in parentheses, * indicates significant at 10%, ** significant at 5%, *** significant at 1%.

Table 3
IV Estimation of augmented Solow model with endogenous diversity indices
(Dependent variable: growth rate GDP per capita;
Instruments used are one-period lagged diversity and latitude)

	(1)	(2)	(3)	(4)	(5)	(6)
Ln I/Y	0.0197	0.0221	0.0257	0.0228	0.0208	0.0227
	(0.0281)	(0.0281)	(0.0281)	(0.0281)	(0.0283)	(0.0279)
ln HK	0.0461***	0.0455***	0.0435***	0.0455***	0.0477***	0.0437***
	(0.0115)	(0.0114)	(0.0112)	(0.0114)	(0.0120)	(0.0105)
ln(n+ g+ δ)	0.00191	0.000989	-0.000647	0.000601	0.00113	0.00360
	(0.00604)	(0.00618)	(0.00626)	(0.00622)	(0.00626)	(0.00606)
Ln Initial Income	0.00498	0.00496	0.00380	0.00473	0.00456	0.00788
	(0.00597)	(0.00595)	(0.00595)	(0.00595)	(0.00638)	(0.00586)
Ethnic Fractionalization	-0.079***					
	(0.0283)					
Ethnic Polarization (α=0.8)		-0.192***				
		(0.0706)				
Ethnic Polarization (α=1.6)			-0.432***			
			(0.137)			
Polarization (MRQ)				-0.059***		
				(0.0211)		
Ethno-linguistic-religious fract.					-0.0106**	
					(0.00416)	
Ethno-linguistic-religious fract. (avg)						-0.00101**
						(0.000488)
Constant	-0.0922	-0.0868	-0.0553	-0.0817	-0.121	-0.103
	(0.107)	(0.107)	(0.106)	(0.107)	(0.113)	(0.105)
Observations	76	76	76	76	76	76
R-squared	0.285	0.277	0.287	0.278	0.248	0.268
			Diagnostics			
Instruments	l(1).eth latitude	l(1).peth08 latitude	l(1).peth16 latitude	l(1).p latitude	l(1).f latitude	l(1).fr latitude
Sargan-Hansen	2.235	2.584	2.896	2.604	3.824	3.0810
	(0.1349)	(0.108)	(0.088)	(0.1066)	(0.0505)	(0.081)
Shea Partial R-sq	0.8793	0.8395	0.7982	0.8246	0.8992	0.99
F-statistic	207.74	130.88	56.74	99.26	107.037	100000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Anderson CCLR	160.70	139.06	121.64	132.31	174.42	696.85
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Pagan-Hall	2.702	2.890	3.208	2.922	1.526	5.259
	(0.2590)	(0.235)	(0.201)	(0.232)	(0.4662)	(0.0721)
RESET	0.64	1.03	0.89	1.03	0.39	1.48
	(0.4235)	(0.3105)	(0.346)	(0.3098)	(0.5312)	(0.2242)

Note: Standard errors in parentheses below coefficients and, in the bottom panel, p-values in parentheses. * indicates significant at 10%, ** significant at 5%, *** significant at 1%.

Table 4

IV Estimation of augmented Solow model with endogenous diversity indices without initial income

(Dependent variable: growth rate GDP per capita; Instruments are lagged diversity and latitude)

	(1)	(2)	(3)	(4)	(5)	(6)
Ln I/Y	0.0243	0.0269	0.0297*	0.0275	0.0255	0.036*
	(0.0169)	(0.0169)	(0.0166)	(0.0168)	(0.0174)	(0.0004)
ln HK	0.0450***	0.0444***	0.0425***	0.0444***	0.0468***	0.048***
	(0.0128)	(0.0128)	(0.0127)	(0.0128)	(0.0132)	(0.013)
ln(n+ g+ δ)	-0.000267	-0.00130	-0.00249	-0.00162	-0.000887	0.00022
	(0.00528)	(0.00534)	(0.00536)	(0.00535)	(0.00544)	(0.005)
Ethnic Fractionalization	-0.0889***					
	(0.0331)					
Ethnic Polarization(α=0.8)		-0.216***				
		(0.0836)				
Ethnic Polarization(α=1.6)			-0.469***			
			(0.169)			
Polarization (MRQ)				-0.067***		
				(0.0254)		
Ethnic-linguistic-religious fract.					-0.0112**	
					(0.00483)	
Ethnic-linguistic-religious fract. (avg)						-0.001
						(0.00043)
Constant	-0.0482	-0.0423	-0.0194	-0.0388	-0.0828	-0.0332
	(0.0569)	(0.0574)	(0.0583)	(0.0575)	(0.0594)	(0.0587)
Observations	76	76	76	76	76	76
R-squared	0.280	0.272	0.284	0.273	0.242	0.251
			Diagnostics			
Instruments	l(1).eth latitude	l(1).peth08 latitude	l(1).peth16 latitude	l(1).p latitude	l(1).f latitude	l(1).fr latitude
Sargan-Hansen (J)	0.403	0.558	0.974	0.616	2.209	1.529
	(0.5257)	(0.4552)	(0.3236)	(0.4325)	(0.1372)	(0.2163)
Shea Partial R-sq	0.8894	0.8538	0.8208	0.8409	0.9105	0.99
First stage F	117.5	84.61	66.8	76.59	149.71	10000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Anderson CCLR	167.33	146.15	130.67	139.72	183.4	684.60
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Pagan-Hall	2.394	2.516	2.760	2.552	1.004	2.809
	(0.3022)	(0.2842)	(0.2515)	(0.2792)	(0.6053)	(0.2455)
RESET	2.11	3.06	2.33	2.88	1.62	2.47
	(0.1465)	(0.0804)	(0.1269)	(0.0896)	(0.2024)	(0.1159)

Note: Standard errors in parentheses below coefficients and, in the bottom panel, p-values in parentheses. * indicates significant at 10%, ** significant at 5%, *** significant at 1%.

Table 5
System GMM Estimation of augmented Solow model with exogenous diversity
(Dependent variable: growth rate GDP per capita)

	(1)	(2)	(3)	(4)	(5)
lagged growth	0.115	0.109	0.118	0.0448	0.00449
	(0.190)	(0.187)	(0.195)	(0.132)	(0.244)
Ln I/Y	0.212	0.216	0.216	0.155*	0.284
	(0.158)	(0.155)	(0.151)	(0.0917)	(0.181)
Ln HK	0.0449	0.0414	0.0329	0.0334	-0.0084
	(0.0594)	(0.0588)	(0.0640)	(0.0428)	(0.105)
Ln(n+ g+ δ)	0.125	0.121	0.122	0.0715	0.127
	(0.120)	(0.115)	(0.116)	(0.0502)	(0.190)
Ln Initial income	0.0553	0.0574	0.0600	0.0480	0.0842
	(0.0630)	(0.0618)	(0.0640)	(0.0439)	(0.135)
Ethnic Fractionalization	0.415				
	(0.606)				
Ethnic Polarization		0.865			
(α=0.8)		(1.257)			
Polarization			0.268		
(MRQ)			(0.387)		
Ethno-linguistic-religious				0.0340	
fractionalization				(0.0262)	
Ethno-linguistic-religious					0.0118
fractionalization (average					(0.0182)
Observations	76	76	76	76	76
Number of countries	26	26	26	26	26
		Diagnostics			
AR(1)	-1.7	-1.03	-1.04	-1.04	-1.01
	(0.284)	(0.301)	(0.299)	(0.159)	(0.313)

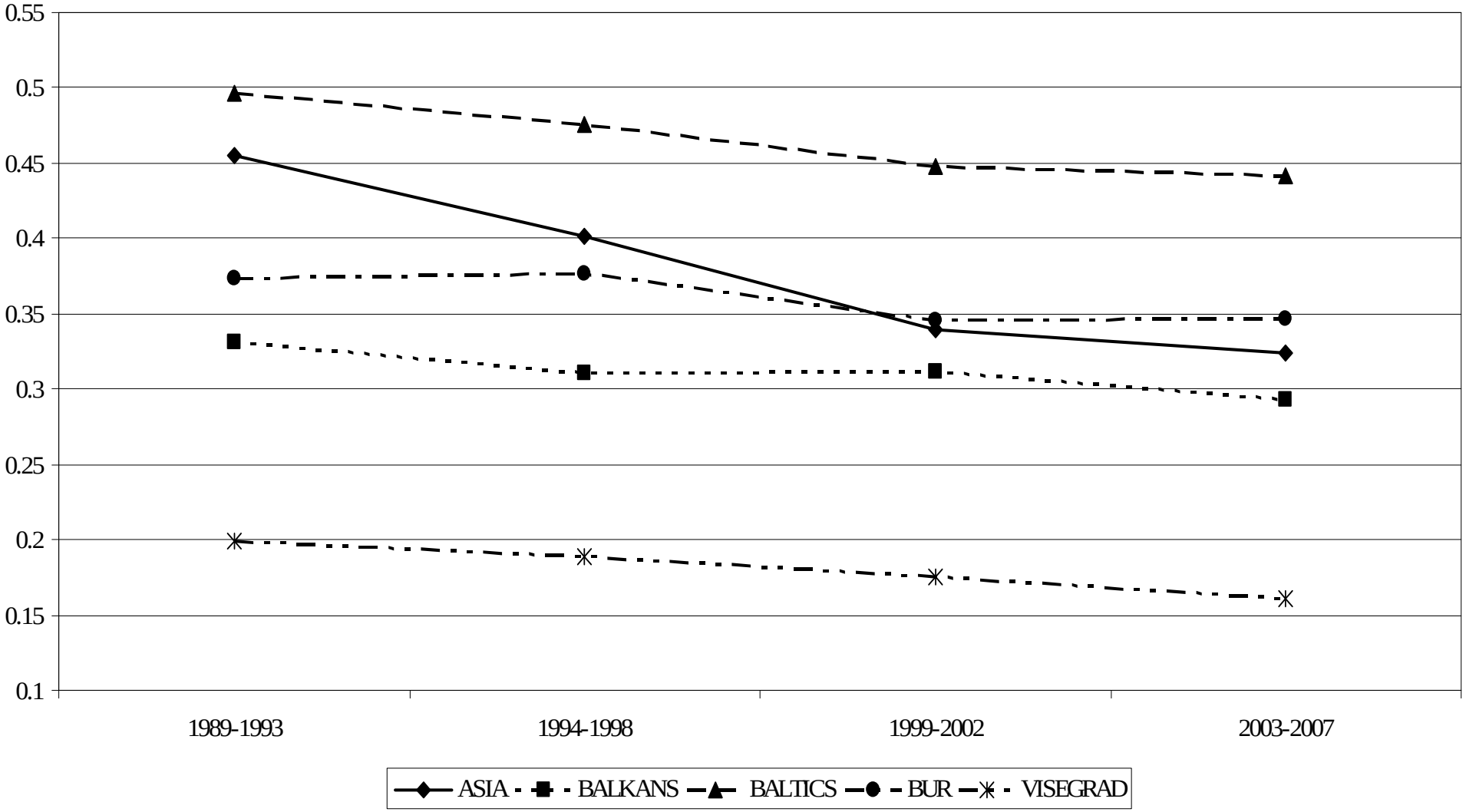
Note: System GMM estimates for growth rate of GDP per capita, 26 transition economies between 1989 and 2007. In the level equation, the instrument used is the first difference of the lagged dependent variable. In the transformed equation, the instrument used is the second lag of the dependent variable. In the top panel, standard errors are in parentheses and * indicates significant at 10%, ** significant at 5%, *** significant at 1%. Period dummies are always included, not reported, and are all significant at 1% in all specifications

Table 6
System GMM Estimation of augmented Solow model with endogenous diversity
(Dependent variable: growth rate GDP per capita)

	(1)	(2)	(3)	(4)	(5)	(6)
L.lng1	-0.114	-0.124	-0.140	-0.123	-0.0846	-0.0868
	(0.105)	(0.102)	(0.100)	(0.101)	(0.112)	(0.126)
Ln I/Y	-0.00240	0.00121	0.0114	0.00350	-0.00646	0.0180
	(0.0210)	(0.0208)	(0.0196)	(0.0202)	(0.0218)	(0.0190)
Ln HK	0.0603***	0.0618***	0.0536***	0.0592***	0.0665**	0.0436**
	(0.0212)	(0.0217)	(0.0180)	(0.0209)	(0.0270)	(0.0196)
Ln(n+ g+ δ)	0.00276	-0.000391	-0.00290	-0.000730	0.00260	0.00654
	(0.00645)	(0.00773)	(0.00736)	(0.00764)	(0.00669)	(0.00743)
Ln Initial income	0.0126	0.0109	0.00665	0.0102	0.0101	0.0178**
	(0.00850)	(0.00858)	(0.00786)	(0.00826)	(0.00866)	(0.00839)
Ethnic Fractionalization	-0.169**					
	(0.0788)					
Ethnic Polarization (α=0.8)		-0.483**				
		(0.216)				
Ethnic Polarization (α=1.6)			-1.036**			
			(0.402)			
Polarization (MRQ)				-0.142**		
				(0.0628)		
Ethno-linguistic-religious fractionalization					-.0206*	
					(0.0116)	
Ethno-linguistic-religious fractionalization (average)						-0.0013**
						(0.000557)
Constant	-0.193	-0.174	-0.0686	-0.151	-0.257*	-0.157
	(0.122)	(0.123)	(0.0991)	(0.117)	(0.142)	(0.122)
Observations	76	76	76	76	76	76
Number of countries	26	26	26	26	26	26
			Diagnostics			
AR(1)	0.91	0.80	0.52	0.73	0.77	-0.08
	(0.363)	(0.423)	(0.604)	(0.464)	(0.441)	(0.940)
Hansen test	19.78	18.69	20.08	20.29	21.29	19.48
	(0.955)	(0.970)	(0.950)	(0.946)	(0.942)	(0.960)

Note: System GMM estimates for growth rate of GDP per capita, 26 transition economies between 1989 and 2007. In the level equation, the instruments used are time dummies and latitude, the first differenced of the dependent variable, investment, human capital, population, and the first difference of the lagged fractionalization index. In the transformed equation, the instruments used are the first difference of latitude, lagged dependent variable, investment, human capital, population and second lag of the respective fractionalization index. In the top panel, standard errors are in parentheses and * indicates significant at 10%, ** significant at 5%, *** significant at 1%. Period dummies are always included, not reported, and are all significant at 1% in all specifications.

Figure 1 Ethnic Fractionalization in Transition: 1989 to 2007



Appendix 1
Correlation matrix

	Log (n+ g+ δ)	Log Human Capital	Log Investment	Ethnic Fractionalization	Ethnic Polarization (α=0.8)	Ethnic Polarization (α=1.6)	Polarization (MRQ)	Ethno-linguistic-religious fractionalization
Log Human Capital	-0.0985							
Log Investment	-0.0906	0.2606						
Ethnic Fractionalization	0.0727	0.0057	-0.0982					
Ethnic Polarization(α=0.8)	-0.0140	0.0042	-0.0477	0.9771				
Ethnic Polarization(α=1.6)	-0.1021	-0.0179	0.0170	0.8629	0.9319			
Polarization (MRQ)	-0.0404	0.0096	-0.0302	0.9619	0.9962	0.9553		
Ethno-linguistic-religious fractionalization	-0.0292	0.0824	-0.0723	0.9174	0.9108	0.8301	0.9028	
Growth rate GPD per capita	-0.1110	0.1981	0.0059	-0.3788	-0.3657	-0.3710	-0.3605	-0.2457