The Road Home: The Role of Ethnicity in the post-Soviet Migration

Youngook Jang
LSE

November 2018

The Road Home: the role of ethnicity in the post-Soviet migration

Youngook Jang

Abstract
This paper argues the importance of ethnic affinity in determining migration patterns using a newly constructed late- and post-Soviet dataset. The members of various indigenous ethnic groups, who had been spread across the Soviet territories, had to decide whether or not to leave the land in which they suddenly became diaspora after the dissolution of the USSR. The migration literature conventionally claims that potential migrants respond to the economic differentials between source and destination, but the post-Soviet case reveals that ethnicity also played a crucial and independent role in migration decision and destination choice. The trend of ethnic un-mixing is evidently seen in the novel dataset regarding the regional migration patterns of major ethnic groups in the post-Soviet space. Econometric analyses using this dataset also confirm that ethnic composition of a region, along with labour market conditions, has significant effects on the regional migration patterns.

Introduction

The conventional wisdom in the migration literature is that potential migrants decide to move when the destination can offer better economic opportunities than the source (Constant and Zimmerman 2013, for survey). However, I argue that ethnic affinity is as important as economic conditions in determining migration decision and destination choice because of the following reasons.

Firstly, as argued by Akerlof and Kranton (2000, 2010), people tend to feel happier and safer when they are surrounded by those who have the same identity as theirs. Ethnicity usually comes up to the surface as an important element of identity when people consider migration. Secondly, migrants would benefit from a specific kind of network effect when they move to the place where their compatriots have already settled, for example, through the channel of access to informal

employment (Frey and Liaw 2005; Kritz et al. 2011). Lastly, policymakers might also want to avoid high ethnic diversity because it is often claimed in the literature that the ethnic fragmentation is harmful to the economic performance, unless there are fine institutions that can deal with the problems arising from this diversity (Alesina and La Ferrara 2005). Therefore, it is possible for the governments to adopt the policies favourable to the majority group and/or discriminative to the minorities. These may affect migration patterns.

While the literature tends to focus on the US and European immigration, the USSR and its successor states afterwards also offer a fruitful ground to see whether this ethnic factor has worked. The Soviet Union was one of the most ethnically diverse political entities in human history, whose residing ethnic groups, or nationalities[1], had been given complex status between autonomy and subjection. Notwithstanding the Marxist ideology of internationalism, Lenin and Stalin committed to national self-determination and established territorial units based on ethnicity (Suny 1989, p.506). 15 of over 100 different ethnic groups were given autonomy, at least in theory, over their own Union Republics (e.g. Russian SFSR, Armenian SSR, Kazakh SSR, etc.) and 20 had their autonomous republics (Tatar ASSR, Yakut ASSR, etc.)[2]. Many ethnic minorities (Poles, Germans, Jews, Koreans, etc.), as well as above-mentioned "titular nationalities" were provided with a certain extent of autonomy including "mother-tongue" education, national units in political parties, and national quotas in colleges, etc. (Slezkine 1994, p.422)[3].

However, the titular groups had not been confined to their autonomous territories but kept being sent, or encouraged to move, towards less developed and less densely-populated regions under the slogan of balanced regional development and

---

[1] Soviet authorities regarded the term "nationality (*Natsional'nost'*, in Russian)" roughly equivalent to ethnicity, and used this as a unit of social accounting, e.g. by recording it in internal passport (*propiska*) or census.

[2] SFSR: Soviet Federative Socialist Republic; SSR: Soviet Socialist Republic; ASSR: Autonomous Soviet Socialist Republic

[3] For more details as to how the ethnic nationality was formed and institutionalised, see Brubaker 1998 and Suny 1993.

"ethnic equalisation". The authorities used first coercions and later incentive mechanisms to control migration of their citizens[4]. This control clearly involved ethnic movements, such as massive forced migration of Volga Germans or Koreans to Central Asia (Polian 2004, p.98) and dispersion of ethnic Russians and Central Asians to Siberia (Kaiser 1994, p.117). These left the members of different ethnic groups distributed across the vast territory of the Soviet Union. Just before the dissolution of the USSR in 1991, the number of Soviet people living outside their own "national" territory was 73 million (Brubaker 1994, p. 57).

The dissolution of the Soviet Union lifted the restrictions on free movement all of sudden. This brought about dramatic shifts in migration patterns. Not only had they become diaspora in the lands where they had long been living, but also they faced the nationalist policies of independent former Soviet Union (FSU, henceforth) republics. The demand for independence of union republics in the late-Soviet period was translated into discriminative policies against non-titular ethnic groups during the state-building process in the 90s (Korobkov 2003)[5]. These formed both a push factor driving the non-indigenous groups away from their living space and a pull factor attracting titular groups to their own national region. Given the preserved ethnic identity and lack of restrictions against the migration among FSU states, the post-Soviet space becomes a perfect stage for researchers to look into the relationship between ethnicity and migration.

Despite this, the previous literature has paid little attention to the ethnic aspects. Most post-Soviet migration studies tend to focus on its relationship with the labour market conditions, insofar as quantitative research is concerned (Mitchneck 1991, Andrienko and Guriev 2004, Guriev and Vakulenko 2015, etc.). The only attempt to quantitatively analyse the ethnic aspect of the post-Soviet migration was Locher (2002). She includes "ethnic sorting" as an explanatory variable to see if migrants tend to move to the areas in which the proportion of their own ethnic group is high.

---

[4] For example, oil and gas industries were established in West Siberia attracting people from the European part (Gibson 1991, pp.147-148).

[5] For example, most Central Asian independent FSU states explicitly discriminated Russian-speaking population in employment or education. For more details, see Korobkov 2003.

She did find the evidence of significant ethnic sorting on 80 ethnic movements during the 90s, though the sample size was small and dataset was not complete.

The previous works have room for improvement in three major areas. Firstly, little attention has been paid to ethnic aspects of the migration movements, though each of the ethnic groups has shown distinctive migration patterns. The ethnic aspect, which may have many interesting implications to the present migration discussion, could not be fully exploited in the previous works. Secondly, none of them has spanned both pre- and post-dissolution periods altogether. The data are concentrating on the post-Soviet period, and some late-Soviet studies did not extend their datasets to the later period. As a result, it was hard to directly compare the different characteristics between pre- and post-collapse migration patterns and their determinants, which may show how institutional changes shape migration movements. Lastly, the analyses on regional migration patterns tend to concentrate on Russian Federation, again mainly due to data availability. Since the inter-republic migration accounts for over 95% of total migrants during the Soviet period and around 80% even after the dissolution, it is crucial to include non-Russian states for a more accurate understanding of Soviet migration patterns.

To improve upon these points, my paper begins with constructing a more reliable dataset on the regional net migration of major ethnic groups using the Soviet and post-Soviet census data and administrative vital records. The newly constructed database will show that many titular ethnic groups which had been distributed across the Soviet territories moved to their own titular FSU states, e.g. Kazakhs to Kazakhstan or Ukrainians to Ukraine, after the collapse. This change in the migration patterns cannot be fully explained by the economic conditions because many have moved from more to less developed regions. Ethnicity must have played a role inferring from the fact that most of the FSU states saw an increase in their titular population after the collapse.

I then move on to use econometric techniques to verify this idea. I regress the constructed regional net migration rates on the share of each ethnic group and

economic conditions, e.g. wages and employment rates, to compare the size of their relative contributions. The results confirm that the trend of ethnic mixing which had occurred in the Soviet period was reversed after its dissolution and the following abolition of migration restrictions. This means that the migrants are now more likely to decide to move out of a region when they have fewer co-ethnics in the region and/or they tend to choose the destinations where the proportion of their co-ethnics in the regional population is high.

All in all, the novel dataset and econometric analysis both confirm the importance of ethnicity in determining the late- and post-Soviet migration patterns. The rest of the paper will proceed to elaborate on the main idea discussed in the introduction. Two following main sections discuss (1) the estimation method and results of the main dataset, the regional net migration of ethnic groups, and (2) the quantitative analyses testing the effect of ethnic affinity in the Soviet and post-Soviet migration. I will conclude in the last section.

Estimating regional net migration of ethnic groups

Data and Method

Though there were some attempts to investigate the extent of ethnic redistribution for the Soviet and post-Soviet periods (Anderson and Silver 1989, Schwartz 1991, Kaiser 1994; Robertson 1996, Heleniak 1997 and 2003), none of these provided comprehensive measures as to regional level migration. This section, therefore, aims to construct the regional net migration of major ethnic groups using the ethnic composition data from censuses and vital statistics published by official statistics agencies of the USSR and FSU states. These agencies provide the relevant data down to the *oblast*/*kray*/ASSR level (*oblast* level, henceforth) as shown in Figure 1.

[Insert Figure 1 here]

The change in ethnic composition over the last decade of the Soviet era can be obtained from the 1979 and 1989 USSR censuses. Since the data are available at the *oblast* level, I can track the population change of a certain ethnic group in an *oblast* over the intercensal period. Provided that the data for the 10-year total natural increase are available, it is possible to calculate the regional level net migration of the ethnic group by the residual method. The estimation strategy to obtain 10-year net migration of ethnic group i in region r can be summarised as follows.

$$M^i_{r,79\_89} = \left(Pop^i_{r,89} - Pop^i_{r,79}\right) - NatInc^i_{r,79\_89}$$
$$\text{where } NatInc^i_{r,79\_89} = w^i_r \cdot s^i_r \cdot Birth_{r,78\_89} - s^i_r \cdot Death_{r,79\_89}$$

$M^i_{r,79\_89}$ is 10-year total net migration, $Pop^i_{r,t}$ is beginning of the year population and $NatInc^i_{r,79\_89}$ is the natural increase from 1979 to 1989[6], which are all of ethnic group i in region r. Using the 10-year crude births ($Birth_{r,78\_89}$), crude deaths ($Death_{r,79\_89}$) of a particular region and the share of each ethnic group in the region's population (s^i_r), I estimate the natural increase of each group in the region ($NatInc^i_{r,79\_89}$). Birth counts are adjusted for ethnic specific fertility weight (w^i_r).

Two publications of *Goskomstat* USSR (1988, 1989) provided the annual data of crude birth and death in 15 union republics from 1979 to 1989, while the regional level natural increase is available for the years 1980, 1985, 1986, and 1988. In order to fill the gap between these years, I assume that the population growth in a union republic was proportionally contributed by the growth rates of population in its *oblasts*[7]. By applying the growth rate of the republic to the regional data of four base years, the annual natural increase for other 6 years can be estimated.

Since all these natural increase data are not available at the ethnicity level, I estimate the natural increase of each ethnic group in a region with the assumption

---

[6] Sum of the natural increase from 1979 to 1988
[7] Rowland does a similar practice when he estimates the natural increase for total population (Rowland 1990, p.659)

that the total natural increase in a region is contributed by the residing ethnic groups according to their shares (s_r^i). That is, if the share of Tatars in the regional population is 35%, then 35% of the natural increase in the region is attributable to Tatars. However, since each ethnic group has different fertility rates, the contribution of each ethnicity to regional birth counts is adjusted for fertility weight (w_r^i)[8]. Using republic level data, I run a regression of the average number of births per woman on urbanisation rates and sex ratio imbalance[9] to predict the fertility rates of each ethnic group in a certain region.

[insert Figure 2 here]

From Figure 2, it is clearly seen that both urbanisation rates and sex ratio imbalance are negatively associated with fertility. In other words, the more urbanised an ethnic group is and the bigger the difference between the numbers of their men and women is, the fewer babies a woman is likely to have in her life. Some regressions confirm this prediction as Column (1) of Table 1 shows that the coefficient of urbanisation is -1.42 and sex imbalance is -2.29, with both being significant at 1% level. I add a dummy variable, Central Asia because it had been reported that Central Asian nationals had higher fertility than others during the Soviet era. Column (2) confirms the prediction by showing the positive and significant coefficient of this variable. Alternatively, in Column (3) and (4), I replace the sex ratio imbalance variable with its logarithmic form, assuming that the relationship between sex imbalance and fertility is non-linear. The regressions with logarithm fit the model better and report a negative coefficient which implies that the curve is concave to the origin. The explanatory power (measured by R^2) is the biggest in the fourth model so I apply the fitted values to each ethnicity in each region to obtain w_r^i, controlling for some outliers[10].

---

[8] The fertility rates of Slavic and Baltic nationalities tend to be lower than those of Central Asian ones.

[9] Sex ratio = | (Male population / Female population) − 1 |

[10] Outliers come out mostly when the number of residents is too small and the ratio of men to women is over 2.

[Insert Table 1 here]

With this weight, the contribution of an ethnic group is inflated or deflated according to the fertility rate of the group. Since I have no data regarding the mortality rates available at the ethnic level and I find little evidence that the rates were significantly different among ethnic groups during the Soviet period, I made an assumption that the death rate is the same across the nationalities. Once the regional natural increase and net migration for each ethnic group are estimated, they can be summed up to union republic and the whole USSR levels.

The same method can be applied to the post-Soviet period. Censuses were conducted in the FSU states around the year 2000[11], and most of them reveal the ethnic information, often at the regional level. The availability and level of birth and death counts vary across the countries, but the biggest states in the area, including Russia and Ukraine, usually have regional level vital statistics. Whereas the *oblast*-level data is sometimes unavailable, namely for Central Asian states, union republic level net migration can be calculated for the whole FSU area for the 1990s.

Results

From the tables and maps below, we can see the regions to and from which the major ethnic groups migrate in the last decade of the Soviet period and the first decade after its dissolution. Tables 2 and 3 summarise the net migration of major ethnic groups at the *raion*-level[12], giving the estimated annual average levels and rates per 1000 in the periods of 1979-89 and 1989-2000, respectively. Figures 3 and 4 shows the annual net migration rates of selected ethnic groups at the level of lower administrative divisions, *oblast*. Red *oblasts* have net out-migration and Blue ones have net in-migration, whereas light yellow means net migration near zero.

---

[11] Only exception is Uzbekistan, where no census has been undertaken since independence. The ethnic composition data was obtained from another source, Ilkhamov and Zhukova (2002).

[12] *Raion* in this paper refers to the economic regions, the basis unit on which the Soviet authority planned economic policies and compare regional performance (Bernard 2013, pp.33-34)

[Insert Table 2 here]

[Insert Table 3 here]

[Insert Figure 3 here]

[Insert Figure 4 here]

On the top left of Figure 3 shows the redistribution of total Soviet population from Central Asia to European part, namely the Baltics and Central *raion* of Russia, and Siberia during the late-Soviet period. This trend is mainly led by ethnic Russians, shown from the net migration of ethnic Russians on the top right. Net out-migration from Central Asia and Trans-Caucasus is seen more clearly, and net in-migration into Slavic and Baltic regions is also quite intense. The out-migration migration of ethnic Russians from non-Slavic southern republics is argued by the previous literature to have begun from the mid-1970s and continued during the transition period (Kaiser 1994, p.166; Zaionchkovskaia 1996, pp. 17-18; Codagnone 1998, p.89; Heleniak 2003, p. 137). As shown in Table 2, the destination of the ethnic Russians arriving in RSFSR seems to have been Northern, Central and Siberian areas. Central *Raion,* including Moscow, was one of the most popular destinations for most ethnic groups except Jews, while Siberia was also attracting the members of many ethnic groups in the 80s.

Similar to the Russians, Ukrainians also show positive net migration in the European part and negative in Central Asia. The difference is that they appear to have left their titular republic and spread across all the regions of Russia and neighbouring republics. This is also seen in Table 2, showing that the Ukrainians out-migrated from Donetsk-Dnieper and South West *raions* of Ukrainian SSR and in-migrated to all other regions, except for Ural, Trans-Caucasus, Kazakhstan and Moldovan *raions*. Two other major titular ethnic groups, Armenians and Kazakhs, show distinctive migration patterns during this period. It is observed that the both ethnic groups have gained population through the net migration in most of the regions of the USSR except in their own titular republics. The figures in Table 2 show that the Armenians and Kazakhs show negative net migration in Armenia and Kazakhstan, whereas it is positive in all other regions.

The net migration patterns dramatically changed after the collapse. The biggest difference is the increase of net out-migration regions, partly explained by the fact that the restrictions on the international emigration were lifted in the late 80s. Also, the absolute number of net migration, which reflects the extent of movement, became much larger compared to pre-dissolution counterparts. The map for all nationalities on the top left shows that the intensifying net-out migration from Central Asia and the new trend of the population losing in Siberia. The European part of Russian Federation appears to gain population through net migration during this period.

Looking at the net migration of ethnic Russians, the fleeing from the southern part of the former USSR territories had continued in an enlarged extent. The out-migration of ethnic Russians was most intense in the war-torn regions, namely Azerbaijan, Georgia, Moldova and Tajikistan where there were ethnic conflicts and civil wars in the late 80s and early 90s. The net out-migration rate of ethnic Russians in Trans-Caucasus *raion* skyrocketed from 24.6 per 1000 in the 80s to 102.5 per 1000 in the 90s, and those in Central Asia and Kazakhstan also marked as high as 70.9 and 50 per 1000, respectively, as shown in Tables 2 and 3. What is different from the 80s is that the ethnic Russians have left not only the southern *raion*s but also all other non-titular FSU states. Their net migration rates in the Ukraine, Belarus, and Baltics were all negative. They appear to have returned to Central and North Caucasus *raions* in Russian Federation, confirming the conclusion of the previous literature with actual data to hand (Codagnone 1998, p.49, Heleniak 2003, pp. 137-139).

[Insert Figure 5 here]

The maps in Figure 5 compare the net migration patterns of each titular nationality in its national republic (including ASSRs), e.g. Latvians in Latvia or Yakuts in Yakut ASSR. It seems that many titular nationalities left their home territories as red regions are more frequently observed, especially in densely populated western parts. It can be inferred that the titular groups were leaving

for other parts of the USSR where their co-ethnics are relatively scarce and thus ethnic mixing was happening. This trend is shown completely reversed in the post-collapse period. The map on the bottom reveals that most of the ethnic groups residing in their titular states mostly have positive net migration rates. This map has more blue regions than its pre-collapse counterpart and this change is more clearly observed in the densely populated European and Central Asian parts. The pattern of return migration was vividly observed in the maps of two major titular nationalities, Ukrainians and Kazakhs in Figure 4, where the net migration is positive only in their homes and negative in all other regions. Ukrainian and Kazakh cases are especially interesting in that the post-Soviet trend had shown dramatic reverse compared to that of the late-Soviet period, as they decided to leave relatively affluent Russia and Baltic countries for less developed Ukraine and Kazakhstan.

The migration trends revealed in my dataset shed light on the motivation of the migrants during the transition period. Firstly, it can be inferred that the ethnicity played a role in population distribution. The changing directions of migration from pre- to post-collapse periods indicate that the trends shifted from ethnic mixing to ethnic un-mixing[13]. Migrants are likely to have considered the share of their own ethnic groups in the destinations when deciding to move, especially after the dissolution. Secondly, the fact that the regions with the highest wages and living conditions, e.g. Tyumen oblast, tend to have positive net migration might suggest that economic conditions also had some influence on migration patterns. This might confirm that income differentials, which have been the conventional determinants of migration, also worked in Soviet and post-Soviet migration decision. Lastly, the regions suffering from civil wars and political conflicts, notably Transcaucasia, show the highest level of net out-migration. It is not too bold to argue that the migrants wanted to avoid conflicts and decided to leave the places. These findings need some verification to be accepted as a solid conclusion, which will be done in the next section with some econometric techniques.

---

[13] Ethnic un-mixing is the process where migration decreases ethnic or national heterogeneity in the receiving regions (Brubaker 1998, p.1047)

Econometric verification

Hypotheses

The economic differentials between source and destination have been conventionally pointed out as a key determinant of migration flows and this hypothesis has been empirically tested for many regions and periods (Constant and Zimmerman 2013, for survey). The role of economic conditions in the Soviet context was also argued to be significant, as discussed in the introduction.

Yet to be fully discussed in the literature is the hypothesis that the ethnic factor is as important as the economic ones in determining migration. I will argue that individuals have an incentive to live in the region where the share of their co-ethnics is high and the government, depending on its institutional arrangements, also prefers its country to be ethnically homogenous.

The starting point is the influence of identity on an individual's economic behaviours. Akerlof and Kranton (2010) define identity as a person's sense of self, and it is affected by his/her gender, ethnicity, religion or any groups they belong to. They argue that a rational individual would seek to achieve their ideal self and may sacrifice monetary compensation for this purpose. Ethnicity, as an element forming identity, would affect migration decision as ethnic identity comes up to the surface when a person migrates to a society dominated by other ethnicities.

In fact, increasingly many scholars have concluded that the ethnic identity of immigrants and labour market outcomes in their destination are closely related (Constant and Zimmermann 2008, 2013, Bisin et al. 2011, Hatton and Leigh 2011, Patacchini and Zenou 2012). If the ethnic identity is one of the most important factors affecting migrants' economic performance, and therefore will affect the decision making and destination selection process of potential migrants in the first place. There is also a handful of research which supports the direct role of ethnicity in the migration decision itself, most of which focus on US ethnic minorities (Kritz and Nogle 1994, Gurak and Kritz 2000, Frey and Liaw 2005). Kritz and Nogle maintain that the larger the number of compatriots residing in a region, the less

likely their out-migration occurs as immigrant communities work as social capital. Similarly, Frey and Liaw suggest cultural constraints such as kinship ties, social support network, and access to informal employment opportunities for the reasons of ethnic sorting happening (Frey and Liaw 2005, p.208).

The ethnic sorting or unsorting through migration is also a matter of serious concern to policymakers, as it is thought to affect the economic development and social stability of a country. A survey paper regarding this matter concludes that ethnically heterogeneous society has been disadvantaged in terms of economic growth while the negative impact may be alleviated when the government can coordinate the diversity (Alesina and La Ferrara 2005, p.763). Poor policy management and political instability are two channels through which ethnic diversity exacerbates economic performance. The conflict of preferences due to the heterogeneity often diverts resources from productive investment to rent-seeking behaviour (Montalvo and Reynal-Querol 2005, p.308). In addition, ethnically diverse societies may be prone to ethnic violence, riots and even military conflicts, as it is more difficult to coordinate the distribution of scarce resources (Vanhanen 1999, p.58). For these reasons, the government may have an incentive to strengthen ethnic homogeneity of the country.

Granted, there has been a view that the diversity can be a good stimulus to the economy under certain circumstances. Lazear (1999, 2000) and Collier (2000) argued that better political institutions can mitigate the adverse effect of ethnic diversity and make the most of spillovers and complementarities arising from different skills, experiences and ideas by coordinating communication. This claim has been supported by some recent empirical works such as Bellini et al. (2013) and Alesina et al. (2016). Therefore, it is not deniable that the government can work to increase the heterogeneity of the country at some stage of its economic development.

In this sense, it is worth noting that the ethnic sorting is not always occurring. Rather, migration is usually associated with the trend of "ethnic mixing", which

happens in the course of attracting foreigners into a society as a policy objective, e.g. in the countries experiencing labour shortage (Brubaker 2008, p. 1047). I will show that ethnic mixing was actually happening before the restrictions on migration were lifted, as a special case observed in the society heavily controlling labour distribution.

Data and Method

I will do the main regressions on the post-collapse dataset (1989-2000[14]), though my dataset spans both late- and post-Soviet periods. Since the two periods show a stark contrast in terms of migration patterns and economic conditions, the pooled estimators do not give meaningful interpretations. Let alone the significant shifts in the political and economic regimes, the abolition of internal passport (or its distributional nature) and emigration restrictions brought a huge difference to the motivation of potential migrants and actual migration decisions. In addition, it is possible to gather more complete regional data of explanatory variables for the latter period[15]. Therefore, I will begin with the regressions on post-Soviet sample as a benchmark, and then will see if the pre-collapse counterpart shows any differences to the former. Where possible, however, all the explanatory variables on the right-hand side are equally included in the specifications of both periods for comparison even though some of these do not have significance in one of two periods.

I select eight major ethnic groups, including ethnic Russians, Ukrainians, Belarussians, Armenians, Kazakhs, Tatars, Jews and Germans to include in my specification[16]. I gathered the dataset regarding the share of major ethnic groups in each region and the performance of different regions so that I can compare their relative impacts on the net migration patterns. As well as having simple OLS estimators, I also try to correct selection bias by two-step Heckman selection

---

[14] The post-collapse sample overlaps about three years of the Soviet period (89-91), it does not affect the discussion too much as the migration restrictions began to be lifted from the late 80s.

[15] Most of the regional level data is available only for 1985 and 90 as to the Soviet period.

[16] The criteria for selection include the size of population, even distribution across the whole USSR, and data availability. The selection issues will be dealt with later.

estimator. The main specification is as follows:

$$(M/P)_{r,t} = a_0 + a_1 EthnicShare_{r,t-1} + a_2 Wage_{r,\,t-1} + a_3 Employment_{r,t-1}$$
$$+ a_3 Conflict_r + (Dummy\_ethnicity)'\beta + X'_{r,t}\gamma + \varepsilon_{r,t}$$

The dependent variable is the net migration per 1,000 persons. Though it would be ideal if there is in- and out-migration data separately, the net migration is also acceptable as an ethnic group usually moved in the same direction in a region, i.e. if they tend to move out of a region, they do so continuously over a certain period. To calculate annual net migration rates, I divide ten-year total net migration by 10 and then by the average population and multiply 1000[17].

I assumed that these regional net migration rates of an ethnic group $(M/P)_{r,t}$ are determined by two major factors, the share of the ethnic group in the region and the economic conditions. The economic conditions of a region can be measured by the level of wages and employment. I included the level and/or growth of wages to measure the compensation which the potential migrants may receive in the source and destination. As there was no unemployment data for the Soviet period – they claimed full employment all the time, – I used the share of employment relative to the regional population in the hope that it can stand proxy for how active the regional economy is[18].

The food consumption level was included to supplement the wage level because Sovietologists often discredit the use of official wage data. In the Soviet period, official labour statistics were prepared by enterprises and organisations on the purpose of administrative reporting. As these enterprises and organisations were evaluated by the authority in terms output and employment targets, they had a strong incentive to distort and falsify the data reported (Clarke 1999, p.273). Also, due to multiple prices (e.g. black market), it is hard to produce real wages from

---

[17] The number of years between two census points is not always 10 for all FSU countries in the post-Soviet period. For example, the first census in Russian Federation was 2002, so I first estimate 13-year total net migration and divide it by 13.

[18] Unemployment rates were available for post-Soviet period, so I use these data where applicable.

the nominal wages. Real consumption is thought to be more accurate in measuring the actual living standard of the Soviet people, so I approximate this real consumption by the level of food consumption per capita in a region.

Besides these main variables, dummy for each of 7 ethnicities except Russians is included to capture the ethnic-specific characters which might have influenced the migration decision. For example, some ethnic groups, such as Armenians showed a tendency to have high in-migration throughout the regions whereas others, notably Jews, showed exactly the opposite. I also included the dummy for war/conflicts, with any war-torn regions being coded 1, and other control variables such as ethnic fractionalisation/polarisation, urbanisation rates, crime rates, number of schools, hospital beds, January/July temperatures etc. in the equation to see if the other factors related to living standards affect the migration patterns[19].

As to the data for the explanatory variables, I mostly consulted the official statistics agencies of the USSR and FSU states to obtain other explanatory variables. The first publication of *Regiony Rossii* (Rostat 2002) contained detailed regional statistics for two years (1985 and 1990) of the Soviet period and the annual data from 1995, and *Narodnoye khozyaystvo SSSR* (*Goskomstat* USSR 1991) and *Official Statistics of CIS states* (CISstat 2008) were utilised for the data of non-Russian republics and Commonwealth of Independent States (CIS, henceforth) states. As three Baltic states were not included in the CIS, I visited the websites of statistical agencies of each state and collected the relevant data.

I first run a regression using a basic OLS estimator. Endogeneity caused by reverse causation is not a big problem in this specification as I used lagged variables for main explanatory variables and it is hard to believe that the future ethnic migration affected the ethnic share or wage levels of the past in any forms. It is especially the case in the Soviet case because the wages were exogenously determined by the government and they were unlikely to be correlated with the post-Soviet counterparts.

---

[19] Definition and sources of all variables can be provided in Appendix if required by referees.

On the other hand, the OLS estimator may be biased because of sample selection problem, as I selected only 8 out of over 120 ethnic groups residing in the Soviet Union. Also, even for the 8 major groups, the ethnic share or net migration data are not available for some regions, especially when the ethnic groups has a negligible share in the regional population. This might cause bias in the coefficient of main regressors because it is hard to believe that the smaller, and thus more-often omitted, ethnic groups tend to have less likelihood of migration.

To avoid this problem, I applied a typical two-step Heckman selection model, with the main selection variable of log regional population in first step probit regression. As summarised in the below equation, I assume that the ethnic data is likely to be missing when its share is below 0.05% of the regional population. The likelihood of missing may be dependent on the absolute size of the regional population, because the bigger the region is, the bigger the population of each ethnic group is, and this absolute size of the ethnic population is often more important than its share for the statistics agency to decide to include this ethnicity in the official data. In other words, even if Kazakhs accounted for about 0.05% in both a small region and a large region, they are likely to appear only in the large region's statistics as the absolute size is bigger. This justification turns out to be quite convincing as will be shown in the result tables later.

$$\left(Ethnicity_{r,t}\right) = b_1 + b_2 \log\left(Population_{r,t}\right) + b_3 Port_r + b_4 \log(Distant\ to\ Moscow)_r + u_{r,t}$$

$$where\ \ I^*\left(Ethnicity_{r,t}\right) = \begin{cases} 1\ \text{if}\ Ethnicshare_{r,t} > 0.05\ (\%) \\ 0\ \text{if}\ Ethnicshare_{r,t} \leq 0.05\ (\%) \end{cases}$$

The use of regional population as the main selection variable also seems to meet the exclusion criterion, as it does not directly affect the migration patterns of an ethnic group. This is because the total regional population is not likely to have an influence on the migration of a certain ethnic group. It would be the population of the ethnic group which matters more than the total population of the region. Moreover, the absolute population does not need to be included as one of the regressors because it is already adjusted when calculating the net migration rates.

In fact, the correlation between total population and ethnic migration rates turns out to be very weak.

I also include the port dummy and distance from Moscow as other variables affecting the selection of net migration data. It is easier to collect the data on ethnicity and include them in the statistics if the region has an access by sea or big river and it is close to the capital.

Results

Table 4 summarises the results from OLS and Heckman 2-step estimations for the post-collapse sample. Along with ethnic share, employment and conflict, I first put wage growth alone in the specification, then replace it with log food consumption, and finally include both.

[Insert Table 4 here]

Looking at the result shown in Column (1), ethnic share in a region turns out to have a positive and significant effect on the net migration of the ethnic group. This implies that a person who has a certain ethnicity tends to move to a place where his or her co-ethnics account for a high share in the regional population, and/or not to move if he or she is already living in such a place. The result remains unaffected by the choice of wage growth or food consumption (Column (2) and (3)). In an econometric sense, one standard deviation change in ethnic share raises the net migration rates by 10.0 to 10.6 per 1000.

It is a confirmation of the hypothesis that the migrants consider the ethnic affinity in a region important when they decide to migrate or choose the destination. As a result of people choosing to move to the places where their co-ethnics live, ethnic un-mixing occurs. At this stage, it is hard to tell whether this effect comes through which of the channels I discussed earlier, namely individual utility maximisation, network effect and policy, it seems clear that ethnicity has an independent influence on migration even after controlling for economic variables.

The economic and conflict variables also have significant coefficients with predicted signs. Lagged wage growth and employment did have a positive impact on regional net migration, reassuring the conclusions of previous studies discussed in the hypotheses section. When replacing wage growth with log of food consumption, it also turns out to be significant and well explains the net migration rates. The relative contributions of food consumption in terms of standardised coefficients, 11.4 - 12.3 rises in net migration rates for 1 SD deviation, seem larger than those of growth (4.2 - 5.2). They stay significant when both are included together, though. On the other hand, the effect of conflict is unclear. Its coefficient is negative as predicted and significant at 1% level in Column (1), but it becomes significant at 5% level when the food consumption is included and eventually insignificant when both wage and food come in the equation (Column (2) and (3)). This may indicate that the effect of food reduction absorbs the effect of war.

Column (4), (5) and (6) report the results from first and second step regressions of the Heckman correction model. About 30 percent of the sample is censored and the selection variables seem to work quite well. The data is more likely to be missing when the absolute population of the region is small, there is no major port and Moscow is far. The significant inverse mills ratio in all three specifications means that there is possible selection bias in the sample which is corrected by the Heckman model.

However, the Heckman estimations do not report too much different results from what we have in the OLS regression. The sign and significance of ethnic share and other economic variables remain unchanged. The relative contributions of ethnic share to net migration rates are larger than the standardised coefficients of wage growth (4.2 - 5.2), but those of food consumption (11.4 - 12.3) and employment share (11.5 - 17.4) turn out to be even larger. The conflict variable, unlike the results from OLS model, still stays significant in Column (6).

[Insert Table 5 here]

I then check the robustness of main variables by trying different combinations of control variables, such as temperature range, average house space, number of students, passenger distance by bus, number of medical personnel and crime rates, etc[20]. Most notable among the results shown in Table 5 is the stability and robustness of the ethnic share variable. It keeps reporting significant and positive coefficients around 0.30. The expected increase in net migration along with one standard deviation change in ethnic share is from 8.3 to 8.6, not too lower than the previous results shown in Table 4. It is safe to conclude that the ethnicity is one of the most important factors affecting the migration patterns of ethnic groups during the decade after the dissolution of the USSR, confirming the hypothesis discussed earlier.

The economic variables also turn out to be quite robust even though their significance does not maintain as stable as that of ethnic share variable. For example, the standardised coefficients for employment share range from 0.17 to 0.37, which results in the net migration rate change by 8.0 to 17.9 per 1000 responding to one standard deviation change in employment. The influence of food consumption is also somewhat sensitive to the changes in specifications, whose standardised contributions vary from 8.4 to 12.8 per 1000. Though the size of effects for these economic variables is not as stable, we still can conclude that the wage growth, employment, and food consumption level are all important in determining net migration rates, given the strong significance of their coefficients throughout.

Another interesting aspect of the result above is that other explanatory variables related to the welfare and living conditions of residents become significant. This is not observed in the late Soviet period and is obviously due to the fact that people now are able to choose the place where they want to move and conventional variables explaining migration patterns begin to work. Living space per capita, education and transport now seem to influence the migration positively, going more in line with conventional prediction. Temperature range between winter and

---

[20] The combination was carefully chosen considering the multicollinearity issue.

summer has turned negative and significant, implying that the bigger the temperature range is in a region, the more likely the migrants leave or the less likely they choose to move there.

Health does not report any significance both in pre- and post-collapse specifications. This is possibly due to the universal health care system, which is regarded as one of a few legacies of Soviet socialism. People had been already benefitting from the relatively equal medical system and there is little reason that they consider health care when they decide to move.

[Insert Table 6 here]

I now run the same regressions separately for each ethnic group to see if there are differences by ethnicity in responding to ethnic affinity and/or economic conditions. In Table 6, you can see that each ethnic group has different responsiveness to the explanatory variables. Russians, Ukrainians and Kazakhs were most sensitive to ethnic share in the destination or the source, showing standardised coefficients of 17.7, 7.0 and 7.8, respectively. On the other hands, Belarusians and Armenians do not seem to care much about the presence of their co-ethnics in the migrating regions. The interesting case is the Jews, who show the negative coefficient on the ethnic variable. This may be because when the proportion of Jewish people is high in a region, they are better organised and help each other to leave the region more easily.

As to the economic variables, each ethnicity shows different sensitivity to them again. Some seem to respond more to wage growth (e.g. Armenians and Jews), and others seem more to food consumption (e.g. Ukrainians and Belarusians[21]). Kazakhs do not consider the economic conditions very seriously as the data constructed in the previous section have suggested.

---

[21] This result is not surprising in that the Slavic countries are famous for their agricultural products.

When I remove possible exception of ethnic Russians and Jews (Column (7) and (8)), the results are reporting the significance of both ethnic and economic variables, which again confirms the robustness of the conclusion I draw previously.

[Insert Table 7 here]

Table 7 summarises the results from OLS and Heckman 2-step estimation for the period before the dissolution of the USSR. The results reveal some significant determinants of late-Soviet migration, but the interpretation is not so simple because migration was still controlled by the Soviet regime during the 1980s. The fact that this government regulation was gradually mitigated undergoing *perestroika* and *glasnost* period in the late 80s makes the interpretation more complex. The significant coefficient could be read as one of the government's priorities in allocating people in the region, or as potential migrants' voluntary motivation.

Column (1) shows that the ethnic mixing clearly occurred during this period, being inferred from the negative and significant coefficients in the OLS specification. Considering the government's population distribution policy, the migrants were likely to be allocated to the places where the share of their co-ethnics was small. This result also implies that if one was living in the region dominated by his or her own ethnic group, he or she is more likely to leave. This is in line with the balanced regional development plan led by the Soviet regime and confirms that the ethnicity was among the key criteria of population reallocation.

The magnitude and significance of the ethnic variable remain unaffected by the changes in specifications and the combination of other explanatory variables. When standardising the size of the ethnic share coefficient, one standard deviation change decreases the net migration rates by about 4.7 to 4.8 per 1000. It seems to have been among the largest and most significant explanatory variables affecting the migration patterns of ethnic groups.

22

The OLS estimation also shows that the wage growth did have an influence on the net migration as well as the employment level. When wage is replaced by food consumption as shown in column (2), it still seems to influence the migration patterns, actually being more significant. Among these three economic variables, employment was the only economic variable which might have contributed to the migration decision relatively consistently. Other things being equal, one standard deviation increase in employment share would raise the net migration into the region by 4.4 to 5.3 per 1000, which is similar to or somewhat higher than the contribution of ethnic share. The significance of wage growth is gone when both the wage and food consumption are included as seen in Column (3).

The results from Heckman correction model are presented in column (4), (5) and (6). The Heckman model does not seem to bring about large changes to the signs and significance of the main coefficients, except the t-statistics for wage variable becomes slightly bigger. We can say that the ethnic share and conflicts negatively affect the regional migration patterns whereas the opportunity of food consumption and employment positively do so.

[Insert Table 8 here]

Lastly, I check whether this conclusion remains robust even after trying different combinations of control variables. Since wage growth becomes insignificant when food consumption is added, I remove wage variable this time. Instead, I add mean house space, proxies for education, health, and transport, share of urban population, temperature, crime rates, ethnic fractionalisation. It seems that employment, house space and education are positively correlated and thus I include them separately. The sign and significance of ethnic share variable survive all the specifications. In fact, all the ethnic share, food consumption and conflict dummy seem robust to the addition of other variables.

Among the control variables, the effects of education and transport are worth noting, as their coefficients are consistently reported to be significant. It is against

the conventional wisdom that the migrants move towards the place providing poorer education and less transportation. The same goes with crime rates, as people choose to stay in the places where there are more crimes. This implies that the Soviet authorities tried to allocate people by artificially manipulating employment level, food consumption or house space in some regions, whose living standards were actually not very high. It seems that other variables such as health or weather had little influence on the migration patterns in the pre-collapse period, and this might be because the migrants had limited autonomy in choosing their destinations.

Conclusion

In this paper, I have tried to find out the key determinants of net migration during the post-Soviet transition period. While the conventional migration literature emphasised the role of economic motivations, I argue that the ethnicity also exerted a significant influence on the migration decision and destination choice.

Using the newly constructed dataset, I conducted OLS and Heckman 2 step estimations on the pre- and post-dissolution periods separately. As the hypotheses predicted, the role of ethnic share turns out to be negative in the late-Soviet period and positive in the post-Soviet period. This result confirms that the trend of ethnic mixing occurring in the Soviet period was reversed after its dissolution and the following abolition of migration restrictions. Economic variables, notably wages, employment and food consumption, seem to explain migration patterns strongly for both the late- and post-Soviet periods. Also, the presence of civil wars did have significant coefficients in some cases but not consistently. The significance becomes weaker when food consumption is included in the post-Soviet specifications and this might indicate that the food shortage was the channel through which the conflicts affected migration. This paper also reveals that different ethnic groups have responded to ethnic and economic factors differently.

Another interesting finding in this paper is that the variables related to the living standards of residents, namely housing space, education or weather, had little influence on migration patterns during the Soviet period, possibly due to the lack of autonomy in choosing destinations. However, the post-Soviet specifications show that the living conditions began to have a stronger impact on migration decision.

I would like to highlight once again the improvement of my method and data compared to the previous literature. First and foremost, it has enabled to include the ethnic aspect in migration analysis. Second, the sample spans both pre- and post-collapse periods, which has never done in the previous quantitative analyses. Last, it captures the effect on long-term migration, which occurred between two census periods usually having a 10-year gap. Though annual data might be useful in observing short-term trend changes more easily, my data can be useful to find out the fundamental determinants of migration. The methods and conclusions suggested in this paper not only contribute to the better understanding of Soviet and post-Soviet migration patterns but also can be used in the analyses of the migration patterns for different regions or periods, as the literature looking at the determinants of migration from the ethnic perspective has been rare. This research combining with future similar studies may enhance our understanding on the recent rise of nationalism and policymakers' intention to increase or decrease ethnic homogeneity in a country by controlling immigration.

Bibliography

Akerlof, G. A., & Kranton, R. (2000). Economics and identity. *Quarterly Journal of Economics*, 715-753.

Akerlof, G. A., & Kranton, R. (2010). *Identity Economics: How Our Identities Shape Our Work, Wages, and Well-Being.* Princeton University Press.

Alesina, A., & Ferrara, E. L. (2005). Ethnic diversity and economic performance. *Journal of economic literature, 43*(3), 762-800.

Alesina, A., Harnoss, J., & Rapoport, H. (2016). Birthplace diversity and economic prosperity. *Journal of Economic Growth*, 21(2), 101-138.

Anderson, B. A., & Silver, B. D. (1989). Demographic sources of the changing ethnic composition of the Soviet Union. *Population and Development Review*, 609-656.

Andrienko, Y., & Guriev, S. (2004). Determinants of interregional mobility in Russia. *Economics of transition, 12*(1), 1-27.

Bellini, E., Ottaviano, G. I., Pinelli, D., & Prarolo, G. (2013). Cultural diversity and economic performance: evidence from European regions in *Geography, institutions and regional economic performance* (pp. 121-141): Springer.

Bernard, P.J., (2013). *Planning in the Soviet Union.* Oxford: Pergamon Press.

Bisin, A., Patacchini, E., Verdier, T., & Zenou, Y. (2011). Ethnic identity and labour market outcomes of immigrants in Europe. *Economic Policy, 26*(65), 57-92.

Brubaker, R. (1994). Nationhood and the national question in the Soviet Union and post-Soviet Eurasia: An institutionalist account. *Theory and society*, 23(1), 47-78.

Brubaker, R. (1998). Migrations of ethnic unmixing in the "New Europe". *The International migration review, 32*(4), 1,047-065.

Codagnone, C. (1998). The new migration in Russia in the 1990s *The New Migration in Europe* (pp. 39-59): Springer.

Collier, P. (2000). Ethnicity, politics and economic performance. *Economics & Politics*, 12(3), 225-245.

Constant, A. F., & Zimmermann, K. F. (2008). Measuring ethnic identity and its impact on economic behavior. *Journal of the European Economic Association, 6*(2-3), 424-433.

Constant, A. F., & Zimmermann, K. F. (2013). *International handbook on the economics of migration*: Edward Elgar Publishing.

CISstat (2008) *Official statistics of the countries of the Commonwealth of Independent States.* Moscow, Russia: Interstate Statistical Committee of the Commonwealth of Independent States

Easterly, W., & Levine, R. (1997). Africa's growth tragedy: policies and ethnic divisions. *The Quarterly Journal of Economics*, 1203-1250.

Frey, W. H., & Liaw, K.-L. (2005). Migration within the United States: Role of Race-Ethnicity [with Comments]. *Brookings-wharton papers on urban affairs*, 207-262.

Goskomstat USSR (1984) *Chislennost' i sostav naseleniya SSSR: Po dannym Vsesoyuznoy perepisi naseleniya 1979 goda* (The size and composition of the population of the USSR: according to All-Union census in 1979), Moscow: *Finansy i statistika*

Goskomstat USSR (1988-89) *Naseleniye SSSR 1987: Statisticheskiy sbornik* (The Population of USSR 1987: Statistical Yearbook), Moscow: *Finansy i statistika*

Goskomstat USSR (1990) *Chislennost' i sostav naseleniya SSSR: Po dannym Vsesoyuznoy perepisi naseleniya 1989 goda* (The size and composition of the population of the USSR: according to All-Union census in 1989), Moscow: *Finansy i statistika*

Goskomstat USSR (1991) *Narodnoe khozyaistvo SSSR v. 1990 goda* (The national economy of the USSR in 1990), Moscow: *Finansy i statistika*

Gibson, J. R. (1991). Interregional migration in the USSR, 1981–1985 and 1971–1975. *The Canadian Geographer / Le Géographe canadien, 35*(2), 143-156.

Gurak, D. T., & Kritz, M. M. (2000). The interstate migration of US immigrants: Individual and contextual determinants. *Social Forces, 78*(3), 1017-1039.

Guriev, S., & Vakulenko, E. (2015). Breaking out of poverty traps: Internal migration and interregional convergence in Russia. *Journal of Comparative Economics.* 43(3), 633-649.

Hatton, T. J., & Leigh, A. (2011). Immigrants assimilate as communities, not just as individuals. *Journal of Population Economics, 24*(2), 389-419.

Heleniak, T. (1997). The changing nationality composition of the Central Asian and Transcaucasian states. *Post-Soviet Geography and Economics, 38*(6), 357-378.

Heleniak, T. (2003). The end of an empire: Migration and the changing nationality composition of the Soviet successor states. *Diasporas and ethnic migrants. Germany, Israel and Post-Soviet successor states in comparative perspective*, 115-140.

Ilkhamov, A. and Zhukova, L., 2002. Etnicheskii Atlas Uzbekistana (Ethnic Atlas of Uzebekistan). *Institut Otkrytoe Obshestvo, Tashkent: Fond Sodeistvia Uzbekistan.*

Kaiser, R. J. (1994). *The Geography of Nationalism in Russia and the USSR*: Princeton University Press Princeton, NJ.

Korobkov, A. (2003). State and nation building policies and the new trends in migration in the former Soviet Union. *The Carl Beck Papers in Russian and East European Studies*, (1702), 78.

Kritz, M. M., & Nogle, J. M. (1994). Nativity concentration and internal migration among the foreign-born. *Demography, 31*(3), 509-524.

Lazear, E. P. (1999). Culture and Language. *Journal of Political Economy*, 107(S6), S95-S126.

Lazear, E. P. (2000). Diversity and immigration in *Issues in the Economics of Immigration* (pp. 117-142): University of Chicago Press.

Locher, L. (2002). Migration in the Soviet successor states. *IZA Discussion Paper* No.602, (also published in *Applied Economics Quarterly*, 48 (1),2002, 67-84)

Mitchneck, B. A. (1991). Geographical and economic determinants of interregional migration in the USSR, 1968-1985. *Soviet Geography*, 32(3), 168-189.

Montalvo, J. G., & Reynal-Querol, M. (2005). Ethnic diversity and economic development. *Journal of Development economics, 76*(2), 293-323.

Patacchini, E., & Zenou, Y. (2012). Ethnic networks and employment outcomes. *Regional Science and Urban Economics, 42*(6), 938-949.

Polian, P. (2004). *Against their will : the history and geography of forced migrations in the USSR*. Budapest ;: CEU Press.

Ravenstein, E. G. (1885). The laws of migration. *Journal of the Statistical Society of London, 48*(2), 167-235.

Robertson, L. R. (1996). The ethnic composition of migration in the former Soviet Union. *Post-Soviet Geography and Economy, 37*(2), 113-128.

Rosstat (1993-2010) *Demograficheskiy yezhegodnik Rossii: Statisticheskiy sbornik* (Demographic Yearbook of Russia: Statistical Handbook). Moscow: *Goskomstat Rossii*

Rosstat (2002) *Regiony Rossii: Osnovnyye kharakteristiki subyektov Rossiyskoy Federatsii* (Regions of Russia: The main characteristics of the Russian Federation). Moscow: *Goskomstat Rossii*

Rowland, R. H. (1990). Economic Region Net Migration Patterns in the USSR: 1979-89. *Soviet geography, 31*(9), 657-678.

Schwartz, L. (1991). USSR Nationality Redistribution by Republic, 1979-1989: From Published Results of the 1989 All-Union Census. *Soviet geography, 32*(4), 209-248.

Slezkine, Y. (1994). The USSR as a communal apartment, or how a socialist state promoted ethnic particularism. *Slavic review, 53*(2), 414-452.

Suny, R. G. (1989). Nationalist and ethnic unrest in the Soviet Union. *World Policy Journal*, 6(3), 503-528. Chicago

Vanhanen, T. (1999). Domestic ethnic conflict and ethnic nepotism: A comparative analysis. *Journal of Peace Research, 36*(1), 55-73.

Zaionchkovskaya, Z. A. (1996). *Migration patterns in the former Soviet Union.* Paper presented at the the Conference Report Cooperation and Conflict in the Former Soviet Union: Implications for Migration, RAND.

Tables and Figures

Table 1. Regression results for average number of births per woman, 1989

Dependent Variable: Average number of births per woman	(1)	(2)	(3)	(4)
Urbanisation rate	-1.43*** (-8.07)	-1.49*** (-8.25)	-1.29*** (-8.75)	-1.35*** (-9.13)
Sex ratio imbalance	-2.29*** (-4.04)	-2.07*** (-3.76)		
Log (Sex ratio imbalance)			-0.20*** (-7.47)	-0.18*** (-6.71)
Central Asia		0.32*** (3.10)		0.19** (2.08)
Constant	3.22*** (36.22)	3.20*** (37.44)	2.37*** (17.55)	2.43*** (17.91)
Total observation	101	101	101	101
R^2	0.62	0.65	0.71	0.73

Note: In parentheses are t-statistics. ***, **, * denote the significance level of 1%, 5%, 10%, respectively.

Table 2. Net migration of the ethnic groups, Annual average and per 1000, 1979-89

Economic Regions	All	Russians	Ukrainians	Armenians	Kazakhs	Tatars	Jews	Germans
North	4,394 (0.7)	7,013 (1.5)	2,060 (7.2)	62 (10.6)	151 (44.2)	565 (13.4)	-202 (-29.1)	-221 (-12.2)
North-West	42,154 (5.3)	36,901 (5.1)	4,833 (23.8)	434 (34.3)	404 (70.8)	434 (8.5)	-3,663 (-28.0)	82 (12.4)
Central	83,169 (2.8)	70,288 (2.5)	10,470 (17.0)	1,998 (36.1)	634 (43.5)	1,560 (6.2)	-5,968 (-23.5)	342 (16.2)
Volga-Vyatka	-25,630 (-3.0)	-11,261 (-1.8)	1,048 (16.2)	106 (28.6)	138 (74.7)	-2,549 (-10.9)	-328 (-20.1)	10 (3.0)
Central Black Earth	-16,093 (-2.1)	-15,844 (-2.1)	365 (1.5)	322 (67.4)	75 (38.5)	100 (14.0)	-309 (-27.7)	47 (14.2)
Volga	-6,746 (-0.4)	-11,524 (-1.0)	20 (0.1)	1,000 (50.7)	1,260 (5.1)	-5,502 (-2.5)	-960 (-22.7)	408 (6.4)
North Caucasia	5,895 (0.4)	-11,881 (-1.1)	4,042 (8.9)	8,576 (28.1)	181 (21.5)	-891 (-14.0)	-1,520 (-36.6)	668 (11.2)
Ural	-57,712 (-2.9)	-27,574 (-1.9)	-3,283 (-7.4)	401 (33.5)	39 (0.2)	-1,914 (-1.0)	-1,174 (-24.6)	-1,424 (-9.5)
West Siberia	86,799 (6.2)	57,317 (4.8)	15,148 (31.2)	787 (67.7)	532 (4.5)	4,879 (13.9)	-479 (-19.0)	-1,884 (-4.6)
East-Siberia	5,335 (0.6)	5,299 (0.7)	2,564 (10.1)	412 (50.5)	290 (37.0)	-1,499 (-12.8)	-423 (-34.6)	-690 (-10.5)
Far East	33,000 (4.5)	23,261 (3.9)	6,888 (12.4)	431 (50.0)	350 (40.8)	-20 (-0.2)	-499 (-21.7)	219 (12.6)
Ukraine[1]	13,808 (0.6)	33,229 (5.1)	-15,994 (-1.1)	1,405 (30.3)	333 (37.7)	-275 (-3.2)	-14,770 (-26.4)	371 (10.3)
West (Baltic)	24,109 (2.8)	16,041 (7.0)	4,601 (21.0)	310 (48.9)	89 (43.0)	224 (14.1)	-895 (-18.9)	-105 (-9.7)
Trans-caucasia	-64,246 (-4.3)	-20,925 (-24.6)	367 (4.2)	-37,509 (-9.9)	222 (67.2)	-1,039 (-29.8)	-1,549 (-34.2)	-113 (-37.8)
Central Asia	-100,262 (-3.4)	-40,024 (-12.1)	2,512 (7.9)	50 (0.6)	5,356 (6.4)	-14,617 (-21.3)	-1,178 (-13.1)	-5,258 (-29.0)
Kazakhstan	-85,835 (-5.5)	-49,184 (-8.1)	-6,662 (-7.4)	292 (17.6)	-12,758 (-2.2)	-876 (-2.7)	-427 (-20.7)	-5,771 (-6.2)
Belarus	-2,823 (-0.3)	16,121 (13.0)	4,549 (17.4)	188 (49.0)	91 (50.3)	145 (12.5)	-2,120 (-17.1)	83 (27.7)
Moldova	-6,392 (-1.5)	2,441 (4.6)	-968 (-1.7)	67 (27.8)	44 (53.3)	34 (11.2)	-1,893 (-26.0)	-525 (-56.1)

Note: per 1,000 rates are in parenthesis. See text for the sources and estimation method.

1) Ukraine includes three economic regions: Donetsk- Dnieper, South-West, and South.

Table 3. Net migration of the ethnic groups, Annual average and per 1000, 1989-2000

Economic Region	All	Russians	Ukrainians	Armenians	Kazakhs	Tatars	Jews	Germans
North	-53,996 (-9.5)	-32,210 (-6.8)	-9,825 (-40.2)	220 (28.7)	-212 (-68.7)	-1,285 (-32.5)	-232 (-52.0)	-333 (-20.9)
North-West	26,474 (3.3)	7,600 (1.1)	-4,582 (-24.0)	1,242 (56.6)	-276 (-46.4)	-171 (-3.3)	-4,871 (-64.4)	117 (15.6)
Central	240,643 (7.9)	164,801 (5.9)	-2,978 (-4.7)	12,613 (89.2)	-277 (-17.3)	2,238 (8.3)	-7,965 (-49.5)	401 (16.3)
Volga-Vyatka	1,396 (0.2)	6,502 (1.1)	-1,173 (-19.0)	1,009 (96.7)	-99 (-50.4)	-355 (-1.6)	-516 (-48.7)	100 (25.5)
Central Black Earth	35,198 (4.6)	33,886 (4.7)	-4,201 (-19.7)	2,122 (109.8)	-73 (-39.0)	450 (43.5)	-374 (-56.4)	307 (57.4)
Volga	60,820 (3.7)	40,714 (3.4)	-6,541 (-21.7)	5,609 (93.3)	1,514 (5.5)	18,457 (7.7)	-1,342 (-48.2)	-1,350 (-23.0)
North Caucasia	178,249 (10.0)	67,554 (5.9)	-11,235 (-27.6)	18,728 (40.0)	-160 (-17.8)	737 (10.8)	-1,545 (-62.8)	-1,967 (-37.7)
Ural	14,881 (0.7)	17,204 (1.2)	-9,094 (-24.1)	2,780 (87.1)	593 (3.3)	-15,432 (-8.2)	-1,528 (-49.0)	-4,238 (-35.1)
West Siberia	7,045 (0.5)	25,942 (2.0)	-11,896 (-23.4)	2,679 (80.7)	121 (0.9)	-1,186 (-3.0)	-881 (-51.7)	-10,100 (-29.0)
East-Siberia	-44,403 (-5.0)	-29,516 (-4.0)	-9,572 (-44.3)	1,098 (62.0)	-372 (-47.9)	-1,773 (-16.6)	-361 (-45.7)	-809 (-13.3)
Far East	-98,778 (-13.5)	-62,574 (-10.6)	-26,096 (-57.8)	515 (35.4)	-478 (-58.3)	-2,786 (-38.6)	-1,011 (-71.3)	-475 (-28.5)
Donetsk-Dnieper	-12,205 (-0.6)	-92,346 (-15.7)	91,846 (6.5)	2,254 (50.4)	-383 (-47.8)	-844 (-16.9)	-8,287 (-79.6)	-281 (-15.9)
South-West	-23,726 (-1.1)	-49,236 (-35.6)	38,622 (2.0)	421 (30.7)	-383 (-47.8)	-305 (-31.2)	-13,341 (-106.0)	60 (6.3)
South	1,437 (0.2)	-31,251 (-12.1)	22,451 (5.5)	1,256 (67.3)	-383 (-47.8)	390 (19.0)	-5,936 (-91.3)	-14 (-1.7)
West (Baltic)	-30,648 (-3.6)	-18,828 (-8.4)	-6,190 (-30.2)	N/A	N/A	4,423 (96.4)	-1,771 (-57.9)	508 (36.9)
Trans-Caucasus	-168,670 (-10.8)	-51,689 (-102.5)	-4,762 (-72.8)	-70,543 (-19.0)	N/A	N/A	N/A	N/A
Central Asia	-303,242 (-8.4)	-143,849 (-52.7)	-18,444 (-70.9)	-2,800 (-32.9)	-13,580 (-13.2)	-29,718 (-55.3)	N/A	-16,944 (-159.8)
Kazakhstan	-322,057 (-20.5)	-208,906 (-39.0)	-36,078 (-50.0)	-557 (-32.9)	129,990 (17.9)	-9,467 (-32.8)	-1,177 (-93.3)	-67,850 (-103.5)
Belarus	11,163 (1.1)	-14,009 (-11.3)	-4,451 (-16.9)	N/A	N/A	-229 (-20.3)	-7,903 (-113.1)	N/A
Moldova	-34,307 (-8.3)	-12,107 (-26.0)	-10,620 (-20.4)	N/A	N/A	N/A	-3,858 (-109.4)	N/A

Note: per 1,000 rates are in parenthesis. The *raion* level net migration data are not available for some ethnic groups mainly due to their negligible size.

Table 4. OLS & Heckman 2-step estimation, after the collapse

Dependent Variable: Net migration per 1,000	(1) OLS	(2) OLS	(3) OLS	(4) Heckman	(5) Heckman	(6) Heckman
Ethnic Share	0.35***	0.33***	0.32***	0.35***	0.32***	0.32***
	(6.07)	(5.77)	(5.76)	(6.23)	(5.87)	(5.88)
Wage	1.27***		0.79**	1.35***		0.90***
	(4.86)		(3.02)	(5.14)		(3.44)
Food consumption		68.67***	62.62***		67.77***	60.43***
		(8.56)	(7.61)		(8.13)	(7.08)
Employment	1.58***	2.27***	2.02***	1.67***	2.30***	2.01***
	(9.24)	(14.75)	(11.53)	(9.50)	(14.69)	(11.41)
Conflict	-23.07***	-9.44*	-10.03*	-24.31***	-12.41**	-13.13**
	(-6.10)	(-2.33)	(-2.49)	(-6.36)	(-3.07)	(-3.27)
Dummy for ethnicity	O	O	O	O	O	O
Selection equation (DV: Ethnic Share over 0.05%)						
Regional population				0.24***	0.24***	0.24***
				(3.93)	(3.93)	(3.93)
Port				0.63***	0.63***	0.63***
				(5.58)	(5.58)	(5.58)
Distance to Moscow				-0.14**	-0.14**	-0.14**
				(-2.90)	(-2.90)	(-2.90)
Total observation	824	824	824	1029	1029	1029
R^2	0.61	0.63	0.64			
Censored obs.				284	284	284
Inverse mills ratio				-18.83**	-12.14*	-12.50*
				(-2.52)	(-1.70)	(-1.76)

Note: In parentheses are t-statistics (z-statistics for Heckman estimators). ***, **, * denote the significance level of 1%, 5%, 10% respectively. For definitions and sources of variables, see the text.

Table 5. Robustness check, Heckman 2-step estimation, after collapse

Dependent Variable: Net migration per 1,000	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Ethnic share	0.30*** (5.54)	0.30*** (5.76)	0.28*** (5.50)	0.29*** (5.68)	0.30*** (5.94)	0.29*** (6.00)	0.29*** (5.88)
Wage growth	1.07*** (3.91)	1.26*** (4.87)	1.20*** (4.70)	0.41 (1.60)	0.50 (1.96)	0.65* (2.54)	0.67** (2.61)
Food consumption	66.68*** (7.29)	45.12*** (4.90)	43.60*** (4.79)	79.07*** (9.82)	92.80*** (10.79)	73.33*** (8.02)	70.19*** (7.54)
Conflict	-12.13** (-2.80)	-16.98*** (-4.07)	-11.51** (-2.73)	-16.44*** (-4.55)	-12.75** (-3.25)	-14.97*** (-3.77)	-13.15** (-3.22)
Employment	2.08*** (10.25)	2.22*** (10.75)	2.03*** (9.76)				
House space				6.81*** (16.33)	7.65*** (16.38)	6.98*** (14.80)	6.60*** (12.84)
Education	0.03 (0.62)	0.14** (3.14)	0.22*** (4.71)		0.19*** (4.39)	0.23*** (5.13)	0.24*** (5.40)
Health	-0.08 (-0.11)	0.58 (0.92)	0.73 (1.17)		0.12 (0.21)	0.65 (1.12)	0.76 (1.30)
Transportation		9.98* (1.98)	14.53** (2.88)			13.80** (2.99)	15.81*** (3.34)
Weather		-1.38*** (-8.29)	-1.31*** (-7.94)			-0.63*** (-3.96)	-0.63*** (-3.98)
Crime rates			0.00 (0.33)				0.00 (0.41)
Ethnic Diversity			-32.16*** (-5.32)				-11.19 (-1.80)
Dummy for ethnicity	O	O	O	O	O	O	O
Selection equation	O	O	O	O	O	O	O
Total observation	1012	1012	1012	1009	1009	1009	1009
Censored obs.	284	284	284	284	284	284	284
Uncensored obs.	728	728	728	725	725	725	725
Inverse mills ratio	-14.11 (-1.92)	-18.16* (-2.32)	-0.47 (-0.06)	-5.19 (-0.72)	-19.76** (-2.70)	-22.99** (-3.04)	-12.99 (-1.79)

Note: In parentheses are z-statistics. ***, **, * denote the significance level of 1%, 5%, 10% respectively.

Table 6. Sensitivity variation by nationality, Heckman 2-step estimation

Dependent Variable: Net migration per 1,000	(1) Rus	(2) Ukr	(3) Bel	(4) Arm	(5) Kaz	(6) Jew	(7) Non-Rus	(8) Non-Jew
Ethnic share	17.7*** (6.12)	7.0** (1.95)	1.8 (1.05)	0.2 (0.04)	7.8** (2.11)	-8.0*** (-3.92)	6.4*** (5.07)	8.1*** (5.37)
Wage growth	7.3*** (4.51)	2.3 (1.09)	2.3 (1.26)	11.0** (2.18)	2.1 (0.52)	6.4*** (3.47)	4.3*** (3.26)	4.8*** (3.69)
Food consumption	5.6*** (2.83)	10.4*** (3.61)	6.1*** (2.65)	10.3* (1.76)	6.7 (1.40)	-2.5 (-0.99)	6.0*** (3.72)	7.5*** (10.80)
Employment	2.8 (1.08)	6.5* (1.91)	16.5*** (7.24)	24.1*** (4.15)	6.1 (1.02)	0.4*** (4.37)	15.1*** (9.10)	13.1*** (8.27)
Control variables	O	O	O	O	O	O	O	O
Dummy for ethnicity	X	X	X	X	X	X	O	O
Selection variables	O	O	O	O	O	O	O	O
Total observation	156	156	156	159	159	156	1098	1099
Censored obs.	49	49	53	54	82	53	398	394
Uncensored obs.	107	107	103	105	77	103	700	705
Inverse mills ratio	-10.35* (-1.73)	-20.98*** (-2.57)	-14.10* (-1.68)	-52.06** (-2.04)	-5.97 (-0.22)	-15.99* (-1.94)	-23.85*** (-3.69)	-23.89*** (-3.84)

Note: Standardised coefficients (Change in net migration per 1,000 responding to 1 SD increase in X) are reported for an easy comparison. In parentheses are z-statistics. ***, **, * denote the significance level of 1%, 5%, 10% respectively.

Table 7. OLS & Heckman 2-step estimation, before collapse

Dependent Variable: Net migration per 1,000	(1) OLS	(2) OLS	(3) OLS	(4) Heckman	(5) Heckman	(6) Heckman
Ethnic Share	-0.21*** (-6.13)	-0.23*** (-6.63)	-0.23*** (-6.65)	-0.21*** (-6.30)	-0.22*** (-6.62)	-0.23*** (-6.66)
Wage Growth	0.30** (2.03)		0.20 (1.34)	0.35** (2.32)		0.27* (1.78)
Food consumption		20.65*** (4.25)	19.52*** (3.96)		16.71*** (3.21)	14.87*** (2.81)
Employment	0.94*** (9.05)	0.97*** (10.58)	0.90*** (8.72)	0.87*** (8.29)	0.94*** (10.13)	0.85*** (8.11)
Conflict	-6.01*** (-2.66)	-6.65*** (-3.10)	-5.79** (-2.58)	-9.16*** (-3.62)	-9.54*** (-4.01)	-8.51*** (-3.46)
Dummy for ethnicity	O	O	O	O	O	O
Selection equation (DV: Ethnic data availability)						
Regional population				0.29*** (4.94)	0.29*** (4.94)	0.29*** (4.94)
Port				0.63*** (5.94)	0.63*** (5.94)	0.63*** (5.94)
Distance to Moscow				-0.11** (-2.42)	-0.11** (-2.42)	-0.11** (-2.42)
Total observation	939	939	813	1,269	1,269	1,269
Censored obs.				348	348	348
R^2	0.50	0.51	0.52			
Inverse mills ratio				-16.59*** (-3.47)	-12.15*** (-2.54)	-12.75*** (-2.65)

Note: In parentheses are t-statistics (z-statistics for Heckman estimators). ***, **, * denote the significance level of 1%, 5%, 10% respectively. For definitions and sources of variables, see the text.

Table 8. Robustness check, Heckman 2-step estimation, before collapse

| Dependent Variable: Net migration per 1,000	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Ethnic Share	-0.22*** (-6.30)	-0.22*** (-6.33)	-0.22*** (-6.51)	-0.21*** (-6.23)	-0.22*** (-6.52)	-0.22*** (-6.58)	-0.22*** (-6.68)
Food consumption	23.40*** (3.38)	28.71*** (3.92)	28.05*** (3.87)	32.98*** (5.02)	35.04*** (5.24)	15.23** (2.23)	20.48*** (2.80)
Conflict	-10.86*** (-4.13)	-12.27*** (-4.82)	-10.22*** (-3.97)	-12.21*** (-4.97)	-10.84*** (-4.20)	-10.42*** (-4.37)	-10.12*** (-4.06)
Employment	0.98*** (8.42)	0.80*** (5.76)	0.56*** (6.75)				
House space				2.65*** (9.53)	1.84*** (4.67)		
Education						-0.27*** (-10.90)	-0.22*** (-5.77)
Health	-0.78 (-1.39)	-1.16* (-1.92)	-0.75 (-1.24)	-0.25 (-0.49)	-0.86 (-1.43)	-0.36 (-0.72)	-0.75 (-1.29)
Transportation	-1.59 (-1.63)	-2.10** (-2.10)	-2.01** (-2.12)	-2.37** (-2.52)	-2.04** (-2.08)	-3.32*** (-3.63)	-2.97*** (-3.14)
Weather		-0.28* (-2.11)	-0.21 (-1.55)		0.04 (0.30)		0.01 (0.04)
Crime rates		0.83** (2.46)	0.87*** (2.62)		0.84*** (2.63)		0.58* (1.78)
Ethnic Diversity			-18.22*** (-4.82)		-12.27*** (-2.90)		-6.71 (-1.48)
Dummy for ethnicity	O	O	O	O	O	O	O
Selection Variables	O	O	O	O	O	O	O
Total observation	1245	1245	1245	1245	1245	1245	1245
Censored obs.	348	348	348	348	348	348	348
Uncensored obs.	897	897	897	897	897	897	897
Inverse mills ratio	-11.02** (-2.16)	-10.28** (-1.94)	-11.76** (-2.22)	-9.45* (-1.89)	-13.32** (-2.52)	-4.2 (-0.86)	-7.58 (-1.45)

Note: In parentheses are z-statistics for Heckman estimators. ***, **, * denote the significance level of 1%, 5%, 10% respectively. For definitions and sources of variables, see the text.

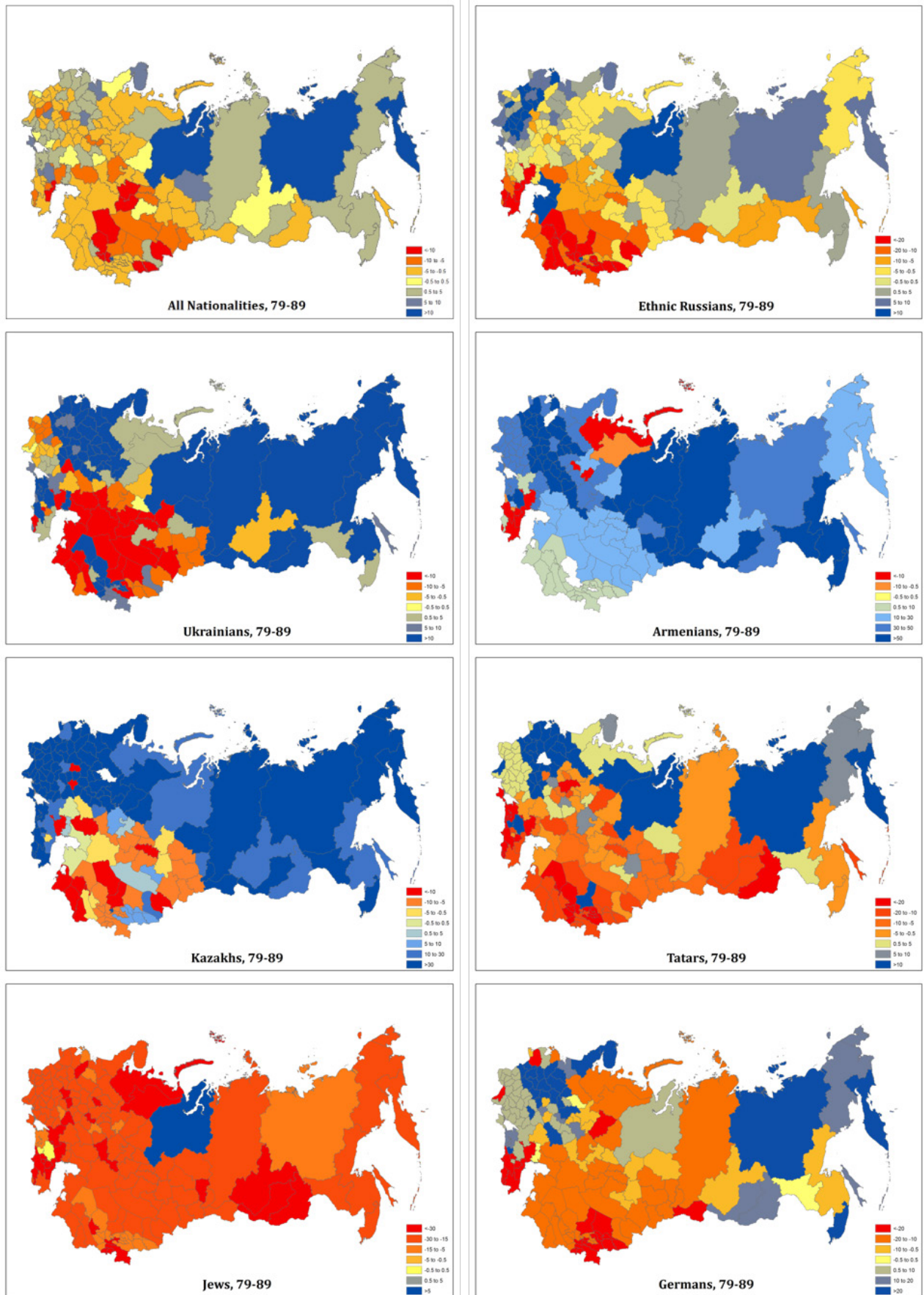**Figure 1. The Administrative divisions of USSR, 1974**

Source: Map Collection, Perry-Castañeda Library, University of Texas-Austin

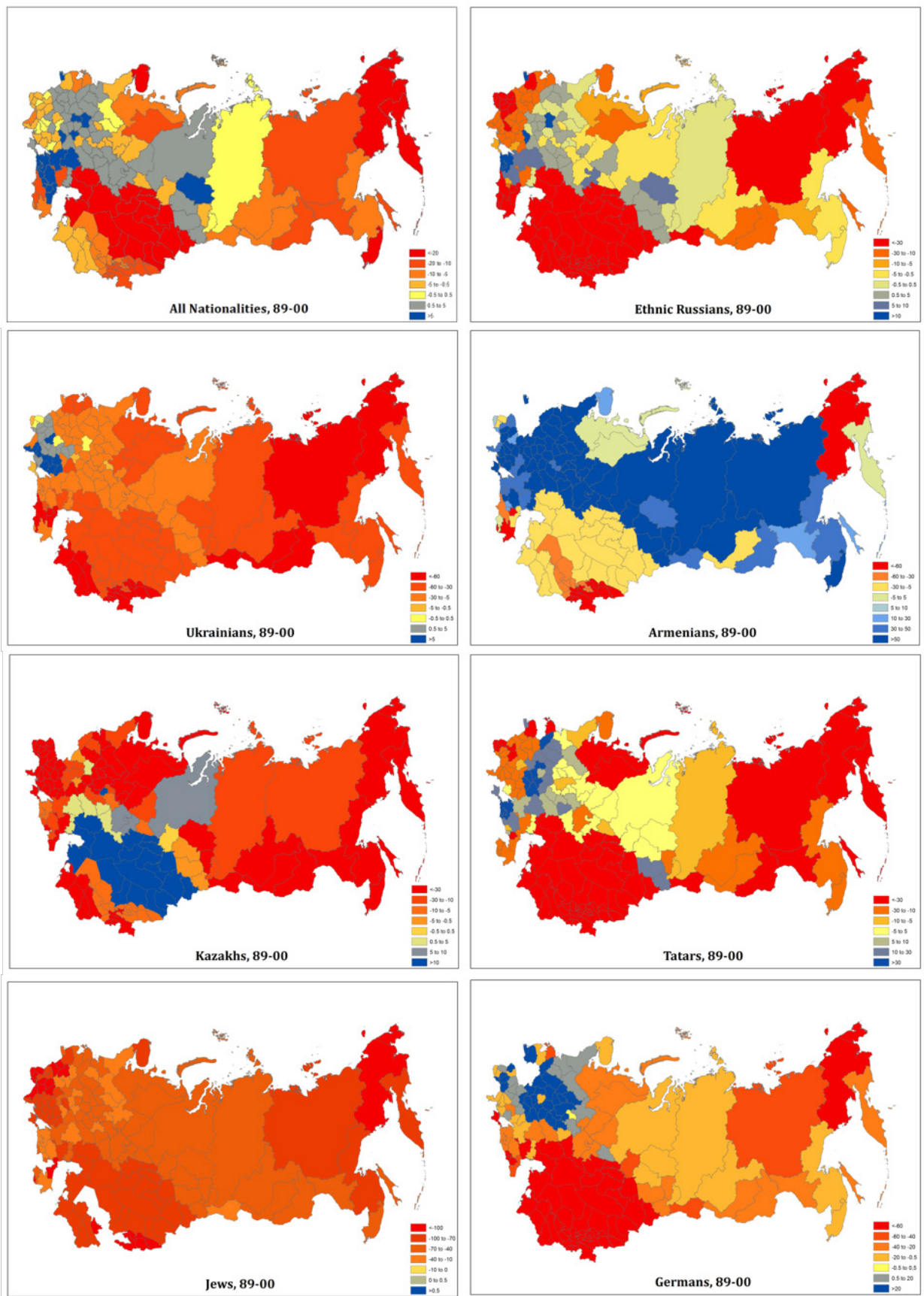**Figure 2. Fertility and urbanisation rates / sex ratios**
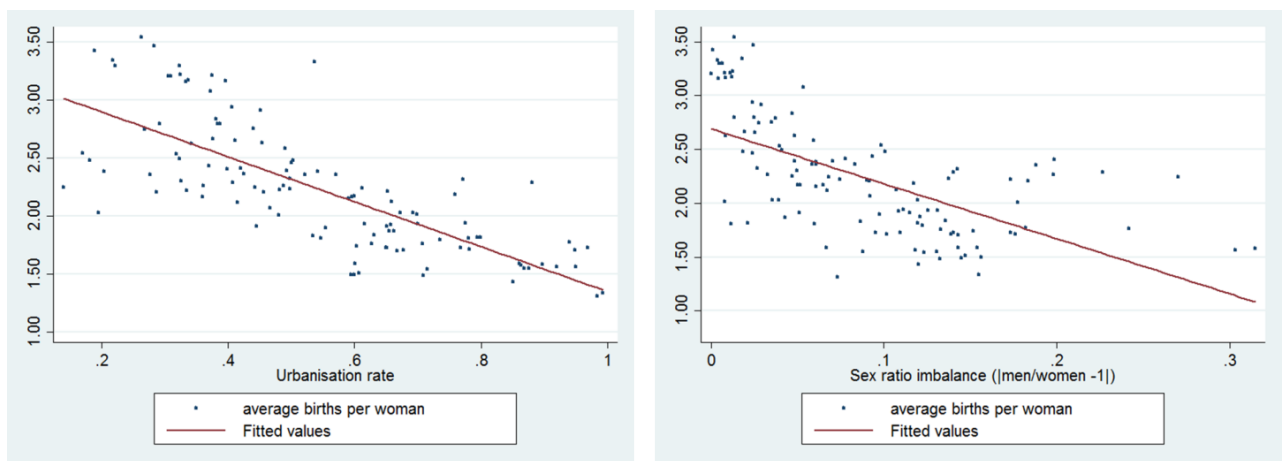
Source: *Goskomstat* USSR (1990

Figure 3. Regional net migration rates of major ethnic groups, 1979-89

Note: Each map has a different legend.

Figure 4. Regional net migration rates of major ethnic groups, 1989-2000

All Nationalities, 89-00

Ethnic Russians, 89-00

Ukrainians, 89-00

Armenians, 89-00

Kazakhs, 89-00

Tatars, 89-00

Jews, 89-00

Germans, 89-00

Note: Each map has a different legend.

Figure 5. Regional net migration of titular ethnic groups, pre- and post-collapse periods comparison

Titular Nationalities, 79-89

Titular Nationalities, 89-00