How the methods of natural sciences can help in the studies of ethnically mixed families?

E Soroko

Institute of Demography at the National Research University Higher School of Economics
20 Myasnitskaya str. 101000 Moscow Russia
Email: esoroko@hse.ru

Abstract. Statistical physics is the branch that uses different mathematical methods in solving not only physical problems. The field of application may be the interdisciplinary studies of many social phenomena. The reason is that they have a stochastic nature. The aim of the paper is to display the opportunities of using the methods of natural sciences in the social sciences. The example is suggested of the research of ethnically mixed families. These are the marital couples where a husband and a wife consider themselves as belonging to different ethnicities. It was demonstrated that application of the reasons used in the kinetic theory helps us to introduce new measure that describes mutual attitudes for a specific combination of ethnicities. The idea of this measure calculation is quite simple. We directly relate the number of marriages established from the reasons of full randomness of collisions of "particles" (persons) and their connection irrespective to their type, and the phenomenology – the actual number of families for a given combination of husband's and wife's ethnicity observed form the population censuses. What we mean by "collision" is any form of personal or social interaction. This measure may be called inter-ethnic propensity, or its inverse value as a mutual inter-ethnic distance.

It was shown that in such multiethnic country like Russia both measures cannot be estimated as the good ones. However this does not mean that the measures introduced are the wrong ones in principle. Simply before their calculation we require to perform co-called "geographical" decomposition that explicitly takes into account the fact and the extent of territorial distribution of population of all the ethnicities in this country by regions. In terms of kinetic approach for gases it may have the analogy of various density of different particles by the volume they are placed in, that is required at consideration of their physical properties.

1. Setting up a problem

Topic of the paper is an interesting and promising interdisciplinary research in demography, sociology, statistics, and ethnography on ethnically mixed families. These are the marital couples where a husband and a wife consider themselves as belonging to different ethnicities. Unfortunately, the standard methods of demographic analysis such as age-sex-specific profiling or life table calculation are not applicable here. Thus it is a rather difficult and curious question "Is 75060 a great deal or a negligible number of Bashkir-Tatar and Tatar-Bashkir families?" (according to the 2010 population census). We cannot answer it simply using some set of appropriate indicators. What are the main reasons? The first one is peculiarity of marriage market formation. On the one hand it is appended by unmarried men

and women depending on their age structure and intensity, on the other hand the numbers of husbands and wives who contract their marriage must correspond to each other. This marriage balance cannot be obtained via usual age-sex indicators. The second reason is that mixed marriage formation is influenced by multiple factors of inter-ethnic interaction. We do not have yet any available indicators for measuring it.

2. Aim of the paper and analogies used

The aim of the paper is to display the opportunities of using the methods of natural sciences in the social sciences. The following analogies may be suggested for better understanding of models for marriage market of mixed marriages.

Table 1. Some analogies in the natural and social sciences used

Natural sciences	Social sciences
Gas kinetic theory	Market of mixed marriages
Particles (molecules)	Bridegrooms, brides (husbands, wives)
Collision	Meeting, conversation, acquaintance, joint work, recreation, education, tourism, journey, sports, etc. In sociology: participation in different small groups
Quantum emission, chemical reaction	Contracting a marriage
Partial pressure of gas A	Proportion of ethnicity A

These analogies give us the opportunity to introduce a set of indicators for estimating the frequency and proportion of the events on the marriage market.

3. How the first approximation may be obtained?

At this step all the absolute numbers and the indicators refer to the country as a whole. Let us introduce the following variables.

M_i – total number of men of ethnicity i in a mixed marriage,

F_j – total number of women of ethnicity j in a mixed marriage,

F – total number of women in a mixed marriage (in the country, equals to sum of F_j),

N_{ij} – number of marital ouples with husband of ethnicity i and wife of ethnicity j (actual, according to the census),

N_{ij}^R – number of marital couples with husband of ethnicity i and wife of ethnicity j (estimated)

Assuming that formation of marital couple (i,j) does not depend on the combination of ethnicities this number may be estimated as

$$N_{ij}^R = M_i * \frac{F_j}{F},$$
(1)

That is formed according to the principle of random equiprobable combination of given ethnicities. From the point of view of kinetic equations of demographic dynamics this assumption was called by Orlov and Suslin "uniform assimilation when there are no preferences at entering a mixed marriage" [1]. It is a result of applying the analogy of random collision of particles in a mixture of several gases. In this case formula (1) may be also interpreted in this approximation as a result of influence of partial pressure of female particles – future spouses of type j on male particles of type i.

Apparently it is not the fact in the general case. That is the number of marital couples depends on this ethnic combination (i,j). For some combinations we observe more mixed marriages and for the other ones – less compared with "uniform assimilation". Let us denote this dependence by multiplier E_{ij}. Then the real number of mixed families (i,j) may be represented in the form

$$N_{ij} = E_{ij} * N_{ij}^R \tag{2}$$

This relation is used for estimating the inter-ethnic preferences E_{ij} :

$$E_{ij} = \frac{N_{ij}}{N_{ij}^R} \tag{3}$$

The approach in this approximation was suggested for the first time in the paper [2] (see in more details).

4. How the matrix of inter-ethnic couples looks like?

Calculations of inter-ethnic indicators are based on the matrix of marriages for all combinations of ethnicities N_{ij} obtained as a demographic table from the census micro-data. New information technology of such census data processing became possible due to the special system developed by Rosstat [3] and described in [4]. From the reasons of visibility a rather restricted set of 11 ethnicities is considered here. In order to provide the demographic balance an additional ethnic group is required. This 12-th conditional ethnic group is suggested to call "the rest". It includes the set of ethnicities not included in the list of selected ones.

Table 2. Matrix of ethnic combinations in marital couples. Russia, 2010, 11 selected ethnicities, number of couples.

Ethnicity of husband	Ethnicity of wife											
	Russians	Tatars	Ukrainians	Bashkirs	Chuvash	Chechens	Armenians	Mordovians	Azerbaijani	Kazakhs	Udmurts	The rest
Russians	12255162	158277	172145	27963	49315	725	14879	39267	3608	11491	28291	262802
Tatars	170031	521326	4402	40381	5477	109	294	1646	292	3178	3026	12404
Ukrainians	212218	5408	58808	1205	1716	25	548	1154	118	537	597	14018
Bashkirs	26484	34679	839	156002	973	7	35	198	49	946	440	2544
Chuvash	47401	5126	1354	1277	148189	3	56	1632	20	170	315	3014
Chechens	4443	245	164	45	16	210093	42	14	37	130	13	2303
Armenians	52546	1042	1557	201	196	22	148271	147	427	103	106	5035
Mordovians	46975	2026	1111	286	2130	11	68	55521	16	123	130	1276
Azerbaijani	29395	2134	1003	394	256	53	662	135	80870	371	185	6032
Kazakhs	16055	3770	552	1296	245	39	41	111	40	77389	72	2531
Udmurts	25391	2379	400	428	242	3	18	66	11	47	51477	1336
The rest	408629	18422	15787	4728	4174	3284	4132	1596	3549	3824	2113	1315065

In this case the matrix of all ethnic combinations of spouses represents a following 12*12 matrix. It contains 15.08 million mono-ethnic marital couples on the diagonal and 2.06 million ethnically mixed couples outside it. Thus, the last ones comprise 12 per cent of all couples and the diagonal – 88 per cent in the Russian Federation in 2010.

Following to the model (1-3) for calculation inter-ethnic preferences we may estimate the distances to the Russians. The results are placed in the following list (ethnicity – distance): Ukrainian – 0.66, Mordovians – 0.68, Udmurts – 0.71, Armenian – 0.74, Chuvash – 0.75,

Tatars – 0.84, Kazakh – 0.97, Azerbaijani – 1.03, Bashkirs – 1.55, Chechens – 2.03 (ethnicities are listed in ascending order of distance).

This matrix may be used for one more new measure in order to describe a propensity to form ethnically mixed marriage with a spouse of any different ethnicity. Numerically it is calculated as a share of ethnically mixed families of a given ethnicity among all the families of this ethnicity. Similar to chemistry, it may be called "valency".

5. The second approximation

More detailed analysis of the results obtained in the first approximation for such multiethnic country like Russia shows that measures introduced cannot be estimated for all the peoples as the good and adequate ones. The reason is a significant inhomogeneity of ethnicity distribution by territory of the country. Some of such peoples have their own national republics, some do not have such administrative-territorial organization but reside in a few number of regions.

Two extreme examples: 98.7 per cent of the Tuvunians reside in their national republic, and 41.2 per cent of the Mordovians. Accordingly, the strong variation for the share of mixed marriages among them is observed. It is 1.9 per cent for the Tuvunians and 49.4 for the Mordovians.

At the second step we require to take into account that ethnic composition of Russia has several significant features that influence its regional peculiarities and the level of inter-ethnic interaction. In addition to the indicators introduced above, some new variables are used for regional population distribution:

M_{iq} – number of men of ethnicity i, in a mixed marriage in a region q,

F_q – total number of women in a mixed marriage in region q,

F_{jq} – total number of women of ethnicity j in a mixed marriage in region q,

N_{ijq} – actual number of marital couples with husband of ethnicity i and wife of ethnicity j in region q,

N_{ijq}^R – number of marital couples with husband of ethnicity i and wife of ethnicity j in region q (estimate),

N_{ij}^{RQ} – number of marital couples with husband of ethnicity i and wife of ethnicity j in the country (estimate based on the regional estimates)

Similarly to (1) in case of full absence of dependence from the specific combination (i,j) a hypothetic number of mixed marriages in region q may be estimated as:

$$N_{ijq}^R = M_{iq} * \frac{F_{jq}}{F_q}, \tag{4}$$

This may be interpreted according to the analogy from the gas kinetics as a result of the effect of local partial pressure of j type females on i type males. In this case the total number of these marriages in Russia may be obtained as a sum of these numbers by regions:

$$N_{ij}^{RQ} = \sum_q N_{ijq}^R, \tag{5}$$

In the general case the estimates calculated from (1) and (4-5) will differ due to significant differences in displacement of the ethnicities by regions of the country.

In order to describe qualitatively the "geographical" factor of displacement of marital couples of ethnicities (i,j) by regions is introduced:

$$G_{ij} = \frac{N_{ij}^{RQ}}{N_{ij}^R},\qquad(6)$$

This variable has the following properties: the value «1» takes place if grooms are uniformly distributed by regions; similarly if brides are uniformly distributed by regions; the values at the diagonal ($i=j$) for monoethnic couples are much more compared with 1 for ethnicities the major part of which resides in the national republic. For instance, according to the 2010 population census this value exceeds 100 for Kabardians, Ossetians, Yakuts, Ingush, Komi, Tuvinians, Carachai.

Similarly to the first approximation we introduce the "new" inter-ethnic preferences E_{ij}^N, free from the factor of population displacement by regions, taken into account as a multiplicative factor:

$$N_{ij} = E_{ij}^N * G_{ij} * N_{ij}^R\qquad(7)$$

Interpretation of this formula is the following. In the absence of influence of inter-ethnic preferences and uniform settlement of ethnicities by territory of the country the number of mixed marriages coincides with the estimate calculated (the first two multipliers are equal to 1). In the absence of influence of inter-ethnic preferences the number of mixed marriages will be distorted by 'geographical' factor of ethnicities settlement compared with the estimate (the first multiplier is equal to 1). If there is no influence of settlement nonuniformity for ethnicities involved (the second multiplier equals to 1) the actual number of mixed marriages for a given combination of ethnicities will differ from the estimated one only due to the influence of inter-ethnic preferences upon these mixed marriages. In the general case "ethnical" and "geographical" factors act combined multiplicatively. Following the relation (7) the estimates of enter-ethnic preferences may be calculated basing on the number of mixed marriages from the population census and the calculated estimates of "geographical" factor from the formula:

$$E_{ij}^N = \frac{N_{ij}}{G_{ij} * N_{ij}^R}\qquad(8)$$

6. "Geographical" decomposition: illustration

As an example for application of the geographical decomposition method for data from the population census the estimated inter-ethnic distances between the Russians and the most common peoples in the Russian Federation may be suggested as illustration in the following table.

Table 5. 32 ethnicities in ascending order of inter-ethnic distance to the Russians. Russia, 2010.

1	Ukrainian	9	Mari	17	Bashkir	25	Avar
2	Mordovian	10	Moldavians	18	Chechen	26	Uzbek
3	Chuvash	11	Armenian	19	Ingush	27	Kabardian
4	German	12	Lezgian	20	Azerbaijanian	28	Tadjik
5	Belorussian	13	Georgian	21	Kumyk	29	Kazakh
6	Udmurt	14	Ossetian	22	Tabasaran	30	Buryat
7	Komi	15	Laks	23	Dargin	31	Karachai
8	Tatar	16	Gipsy	24	Yakut	32	Tuvinian

7. Problems not solved yet

As a new problem statement in ethnography not solved yet an analogy with thermodynamics is suggested for analysis of ethnical population structure and its evolution. Some questions in this field are: Is the entropy actually growing over time as applied to the composition of population by ethnicities? May the dynamics of the population of the USA considered as the well-known "melting pot" for ethnicities be interpreted in the way similar to the second law of thermodynamics? Why this law is not valid in the general case for population ethnic structure at the level of city or country?

8. Conclusion

Using of methods from natural sciences lets us produce much more clear explanation, more simple understanding, modeling, interpretation of the processes under consideration. Models for estimating the inter-ethnic distances give a set of new quantitative measures in ethnography and demography for important processes at the borders of ethnicities – mixed marriages.

9. References

[1] Орлов Ю Н и Суслин В М 2003 Кинетические уравнения для некоторых моделей демографии *Матем. Модел.* **15 № 3**, сс 43-54 [Orlov Yu N and Suslin V M 2003 Kinetic equations for some demographic models *Math. Model.* **15 No 3**, pp 43-54 in Russian]

[2] Сороко Е 2014 Этнически смешанные супружеские пары в Российской Федерации *Дем Обоз* **4** сс 96–123 [Soroko E 2014 Ethnically mixed married couples in the Russian Federation *Dem Rev* **4** pp 96-123, in Russian, http://demreview.hse.ru]

[3] Росстат 2012 Всероссийские переписи населения 2002 и 2010 годов. SuperWEB2. [Rosstat 2012 The 2002 and 2010 All-Russian population censuses URL: http://std.gmcrosstata.ru/webapi/jsf/tableView/customiseTable.xhtml]

[4] Soroko E and Voronin V 2014 New information technology in data processing of population census results *Innovative Information Technologies: Materials of Int sci-prac conf* **Part 3** ed S U Uvaysov pp 430-434