Extending the General Growth Balance Method to Account for Migration

Kirill Andreev

United Nations Population Division

UNITED NATIONS COORDINATION MEETING ON THE ESTIMATION OF ADULT MORTALITY
United Nations Secretariat Department of Economic and Social Affairs
Population Division
New York, 31 July 2008

Objectives:

1) to investigate performance of the General Growth Balance method (GGB) in populations with relatively high migration levels

2) to evaluate current methods developed for adjusting for migration

- How important is it to take migration into account?
- What bias into mortality estimates is introduced by ignoring migration?
- Is it possible to correct for migration in the GGB method?

General Growth Balance Method

Basis: The Balancing Equation of Population Change

$$P_2 = P_1 + B - D + G$$

Assumptions:

a) population is closed to migration, G=0;

b) completeness of first census, k_1, is independent of age;

c) completeness of second census, k_2, is independent of age;

d) completeness of intercensal deaths, c, is independent both of age and year;

GGB regression: $$b(x+) - r(x+) = \beta_0 + \beta_1 d(x+)$$

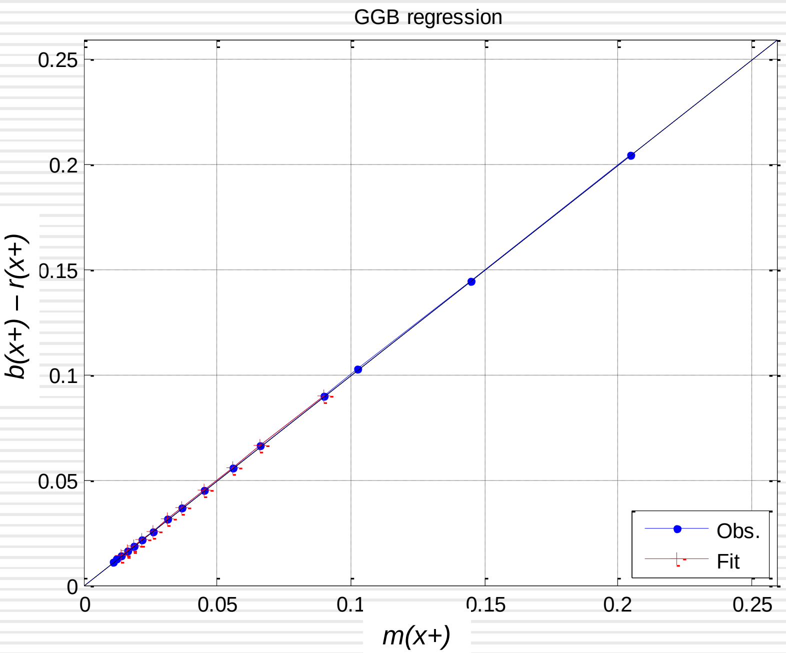
Relative completeness of censuses $$\frac{k_1}{k_2} = \exp(t\hat{\beta}_0)$$

Completeness of death registration $$c = \frac{(k_1 k_2)^{\frac{1}{2}}}{\hat{\beta}_1}$$

Adjusting observed death rates $$m = \hat{\beta}_1 m^*$$

r(x+) – population growth rate above x
b(x+) – entry rate at age x+, “birth rate”
d(x+) – open age death rates
*m** – observed death rates
m – adjusted death rates

Ideal case:



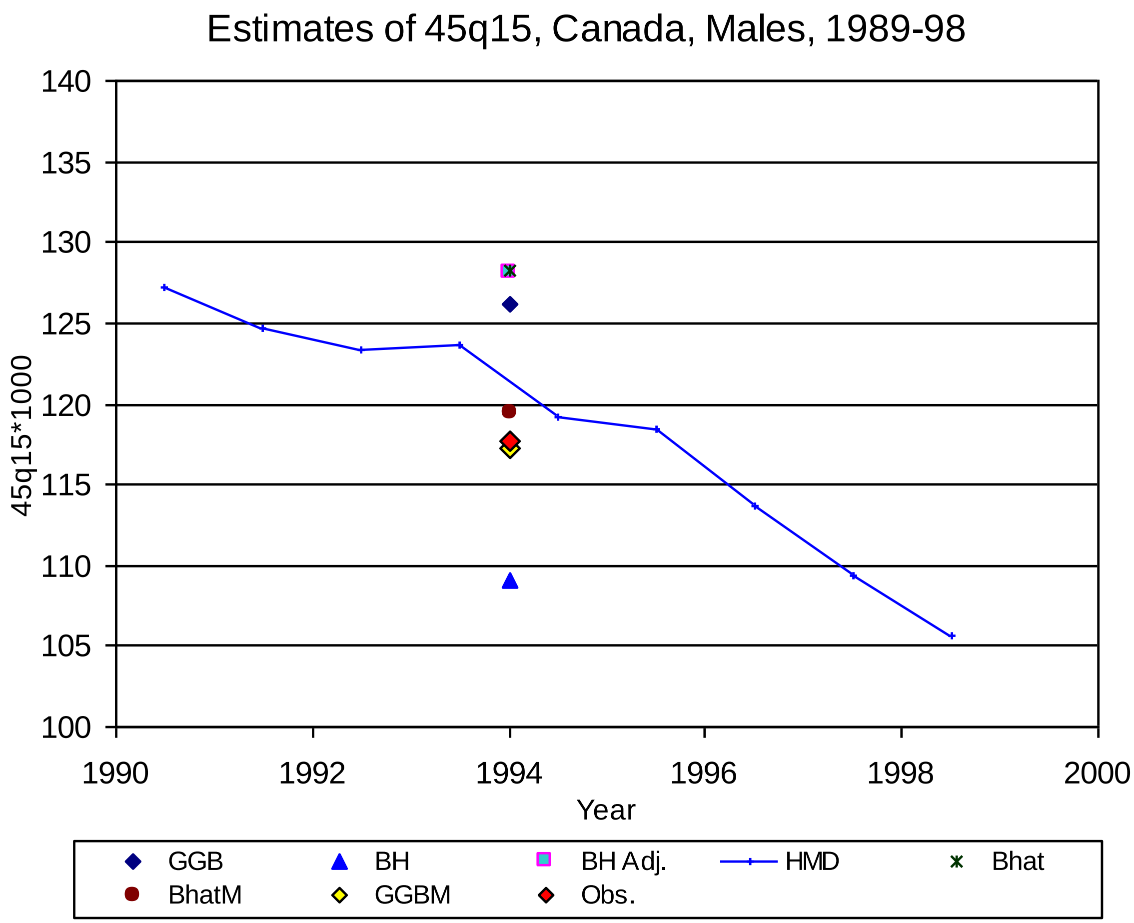
Intercept = 0
Slope = 1

GGB regression with migration:

$$b(x+) - r(x+) + \underline{g(x+)} = \beta_0 + \beta_1 d(x+)$$

g(x+) – net migration rate, age x+

Canada, Males, 1989-1998: effect of unaccounted migration on adjusted death rates and on the relative completeness of death registration



Adjusted 45q15 $m = \hat{\beta}_1 m^*$

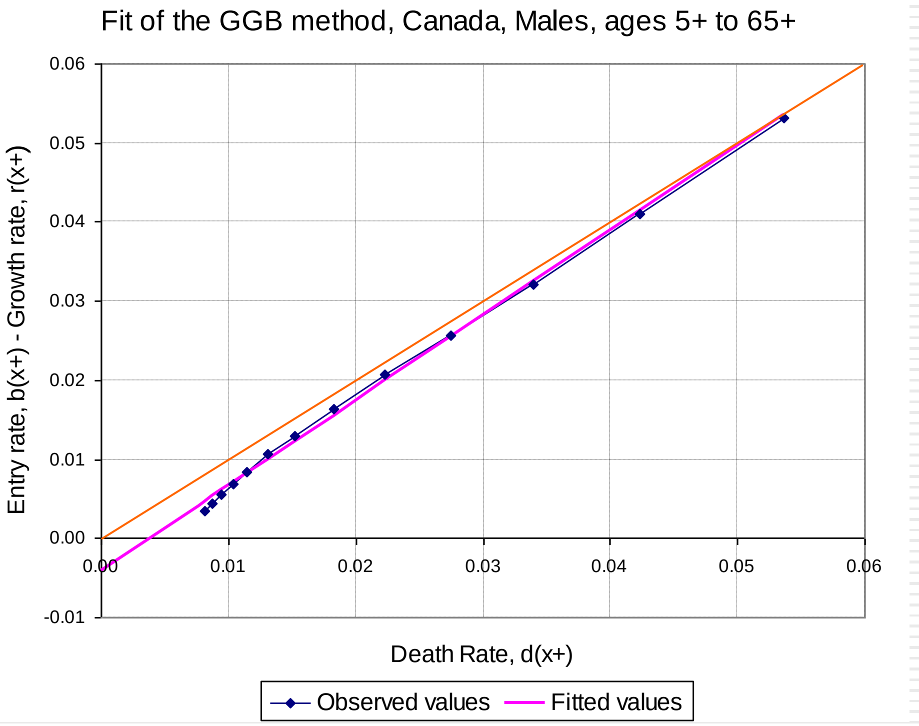
	Level	Rel. Diff.
GGB	126	7.2%
BH	109	-7.4%
BH Adj.	128	8.7%
Bhat	128	9.0%
BhatM	120	1.5%
GGBM	117	-0.4%
Obs.	118	

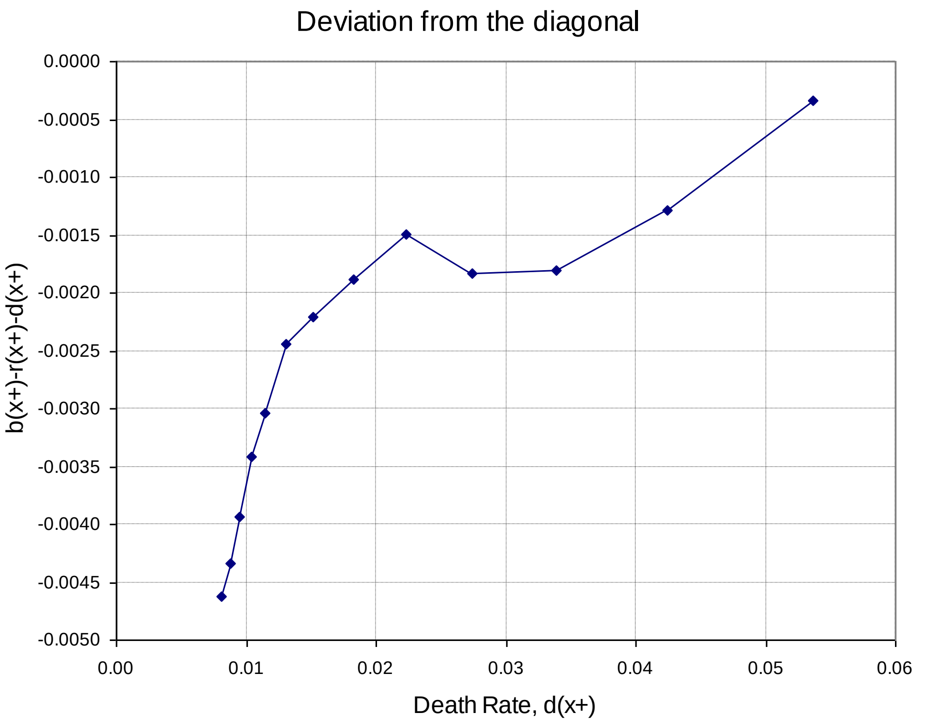
Legend

No migration	
GGB	General Growth Balance method
BH	Bennett-Horiuchi
BH Adj	Bennett-Horiuchi with GGB adjustment
Bhat	Bhat (2002)
HMD	45q15 from annual HMD life tables
Obs	Observed 45q15
With migration	
GGBM	GGB with migration term
BhatM	Bhat (2004) adjustment for migration

- Independent estimates of migration for 1989-1998 provided by Statistics Canada
- Data are considered to be complete with accurate age reporting

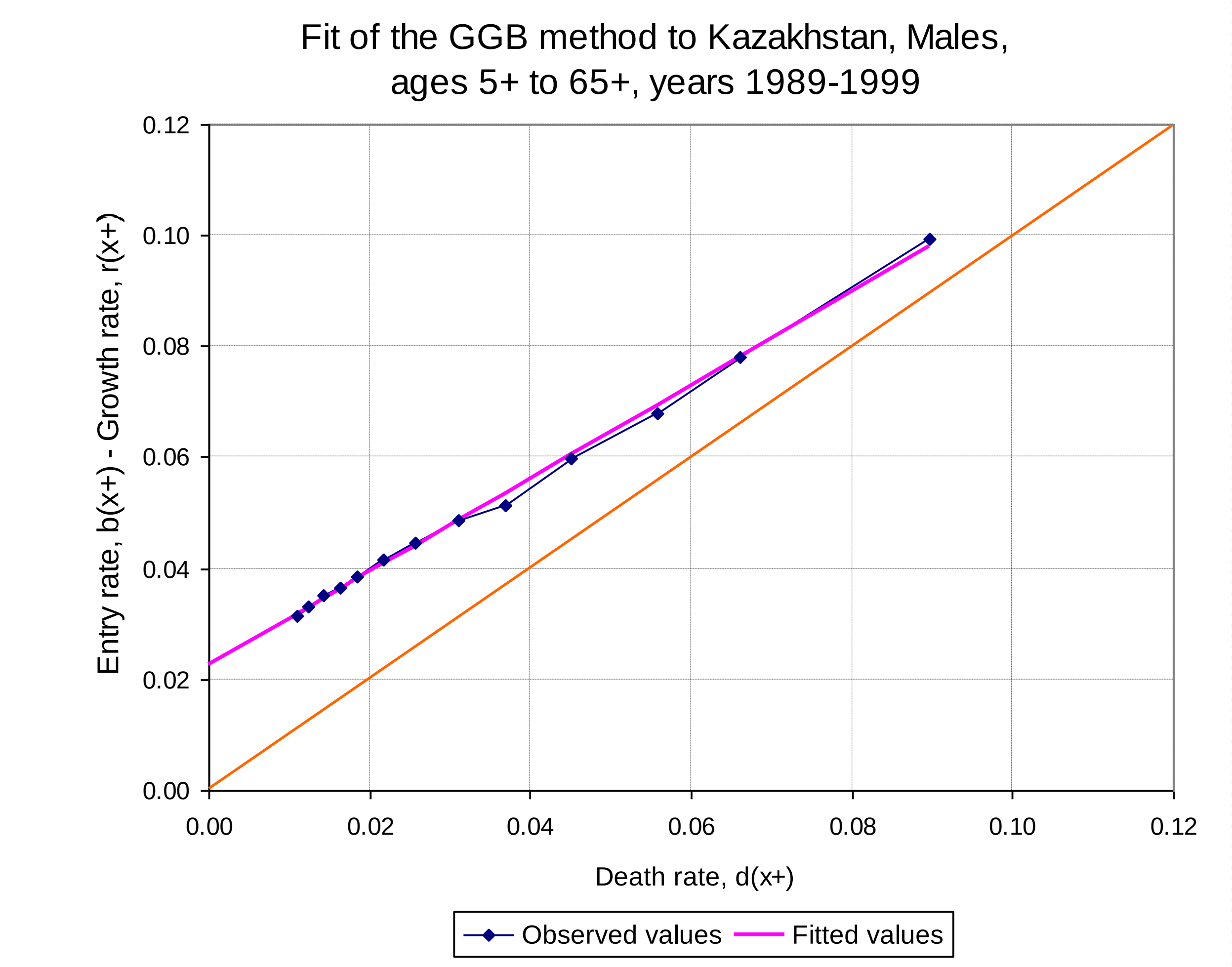
Canada, Males, 1989-1998: diagnostic plots of the GGB method





Slope (adjustment factor) = 1.072 or 7.2%

Kazakhstan, Males, 1989-1999: a case of significant out-migration



Slope (adjustment factor) = 0.84 or 16% reduction of observed 45q15 is need

WHO estimates of completeness and coverage (Mathers, 2005):

Member State	Years of mortality data (all causes)	Completeness[b]	Years with cause-of-death data	ICD revision used[c]	Coverage[d]	Deaths coded to ill-defined codes (%)[e]	Quality[i]
Kazakhstan	1981–2001	89	1981–2001[a]	9	80	5[g]	Medium

Accounting for migration: how to select *g(x+)*?

GGB:

Regression:

$$b(x+) - r(x+) = \beta_0 + \beta_1 d(x+)$$

Objective function to be minimized:

$$f = \left[b(x+) - r(x+) - \beta_0 - \beta_1 d(x+)\right]^2$$

GGB with migration correction:

Regression:

$$b(x+) - r(x+) + g(x+) = \beta_0 + \beta_1 d(x+)$$

Objective function to be minimized:

$$f = \left[b(x+) - r(x+) + g(x+) - \beta_0 - \beta_1 d(x+)\right]^2$$

Approaches to account for migration:

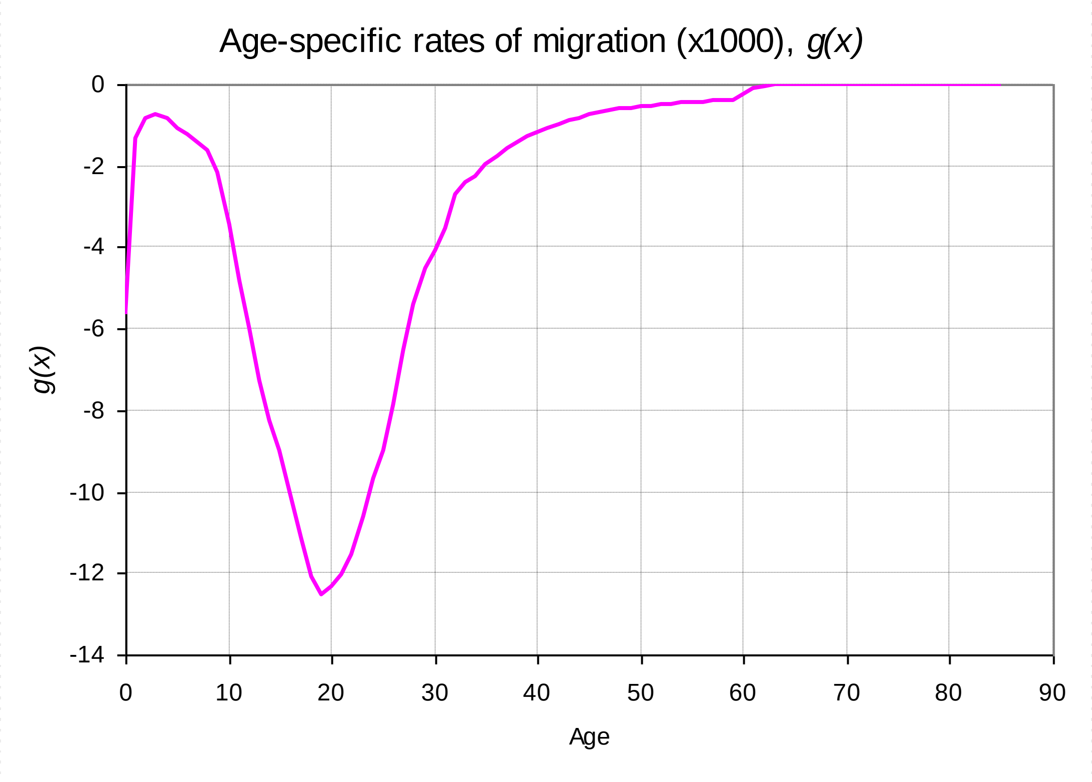
a) Use external $g(x+)$ estimates;

b) Select ages with $g(x+) \approx 0$ to fit the GGB regression, e.g. > 30 or 40;

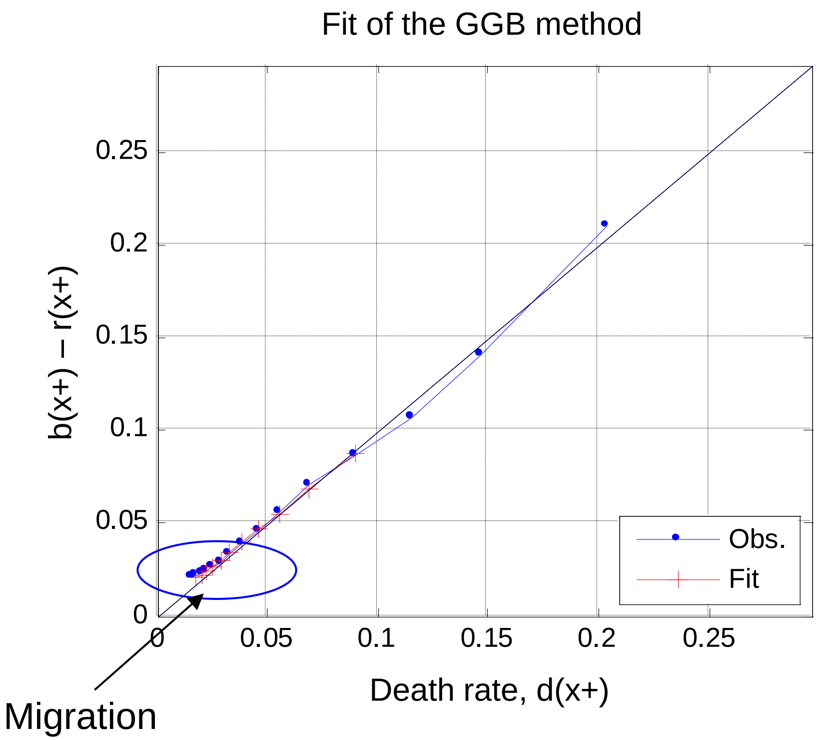
c) Use parametric model for $g(x+)$ e.g. Rogers-Castro migration model.

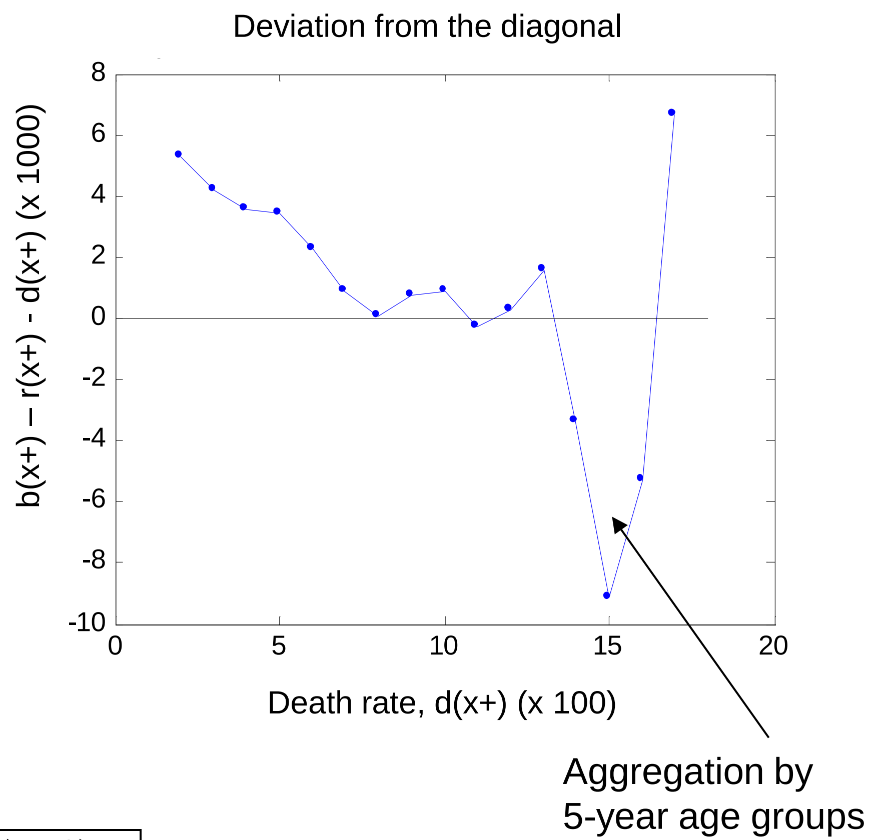
Simulation study (DnkA):

1) Population projection, 20 years, by single year and age (software RUP, USCB);
2) Age structure: Denmark, 1880-1900, Males;
3) Mortality: constant over time, life expectancy at birth = 44;
4) Fertility: constant over time; TFR = 4;
5) Migration: labor out-migration based on residual asdjusted migration pattern



Fit of the GGB method to the simulated dataset





Estimates\Dataset	DnkA	DnkA (g(x) ≡ 0)
Slope	0.9375	0.9833
Intercept	0.0035	0.0004
k1 / k2	1.0677	1.0081
c	1.0323	1.0128
Adult mortality, 45q15		
Observed	0.391	0.390
Adjusted	0.371	0.385
Simulated	**0.394**	**0.394**

Parametric models and identifiability problems

Objective function to be minimized:

$$f = \left[b(x+) - r(x+) + g(x+) - \beta_0 - \beta_1 d(x+) \right]^2$$

Examples of non-identifiable models:

1) *g(x)* = constant: age distribution of migrants is close to age distribution of population e.g. Kazakhstan, period 1989-1999;

2) shape of *g(x)* is close to shape of death rates *d(x)* (unlikely);

3) $g(x+) = g_1 + g_2 g^s(x+)$ where $g^s(x+)$ is a standard pattern based on Rogers-Castro model (Hill and Queiroz, 2004)

Identifiable parametric model

$$g(x) = g_1 g_s(x)$$

$$f = \left[b(x+) - r(x+) + g_1 g_s(x+) - \beta_0 - \beta_1 d(x+)\right]^2$$

g_1 – intensity of migration rates

$g_s(x)$ – as a special case of Rogers-Castro migration model

$$\begin{aligned} g(x) &= a_1 \exp(-\alpha_1 x) \\ &+ a_2 \exp\{-\alpha_2 (x-\mu_2) - \exp[-\lambda_2 (x-\mu_2)]\} \\ &+ a_3 \exp\{-\alpha_3 (x-\mu_3) - \exp[-\lambda_3 (x-\mu_3)]\} \\ &+ c \end{aligned}$$

with child and adult components only and with $a_1=0.02$; $\alpha_1= 0.1$; $a_2=0.06$; $\mu_2=20$; $\alpha_2= 0.1$; $\lambda_2=0.4$ and the rest of parameters set to zero

Fit to simulated data

Estimates\Dataset	DnkA (as before)	DnkA1 (50% underreporting of deaths)	DnkA2 (50% underreporting of population in the second census)
Slope	1.0106	2.0215	0.7146
Intercept	-0.0023	-0.0021	0.0342
k1 / k2	0.9576	0.9574	**1.9152**
c	0.97	**0.48**	1.01
Observed	0.391	0.219	0.504
Adjusted	**0.394**	**0.394**	**0.394**
Simulated	**0.394**	**0.394**	**0.394**

Conclusions and future plans

- If the shape of migration rates is significantly different from the shape of death rates, it is possible to adjust for migration given that age reporting is acceptable and assumptions underlying the GGB method are satisfied;

- A decision, as to whether a particular dataset satisfies the above conditions, should be made beforehand, by exploring raw data and by analyzing external information. Simulation studies could be very useful in exploring distortions introduced in the ideal GGB line by various data problems.

- GGB adjustment should be applied cautiously to the countries with significant emigration as it introduces a downward adjustment into observed death rates;

- We plan to apply Castro four-parameter model of net migration (United Nations, WPP, 1988)

- We plan to apply the GGB method with migration corrections to Central Asian republics of Kazakhstan, Kyrgyzstan, Tajikistan, Turkmenistan, and Uzbekistan.