Session 1: Innovations in the Population Projections at the UN Population Division: Methods, advances and challenges

Probabilistic Projections of the Total Fertility Rate for All Countries: an introduction to the new 2010 UN fertility projection model

Funded by NICHD grant number 1 R01 HD054511 01 A1

Leontine Alkema¹, Adrian Raftery², Patrick Gerland³, Sam Clark², Francois Pelletier³, Thomas Buettner ³, Gerhard Heilig³

¹ National University of Singapore, Singapore, ² University of Washington, Seattle, ³ United Nations Population Division, New York

International seminar on Population Estimates and Projections: methodologies, innovations and estimation of target population applied to public policies
Latin American Population Association (ALAP) - Research Network on Population Estimates and Projections, and National School of Statistics (ENCE) of the Brazilian Institute of Geography and Statistics (IBGE)
Rio de Janeiro (Brazil), 9 -11 November 2011
TFR time series since 1950 can be described with 3 phases:
1. Pre-transition high fertility
2. Fertility transition
3. Post-transition low fertility

We modeled the 5-year changes in the TFR in Phase II and III, using UN estimates 1950-2010.

The observation period is split into the different phases:
Start of Phase II is before 1950 (if max TFR is below 5.5 children), or at latest local max. within 0.5 child of global max.

- All countries are currently in phase II or III.
**Start of Phase III**

Formal definition is based on model parameters (later in presentation)

Within observation period:
Start of phase III is approximated by the midpoint of earliest two subsequent increases below 2

Start of phase III before 2005-2010 observed in 21 countries
(Belgium, Bulgaria, Channel Islands, Czech Republic, Denmark, Estonia, Finland, France, Germany, Ireland, Italy, Latvia, Luxembourg, Netherlands, Norway, Russian Federation, Singapore, Spain, Sweden, United Kingdom, United States of America)
Outline

1. Phase II: The fertility transition
2. Phase III: Post-transition low fertility
3. Results
**TF projection for high fertility countries**

Deterministic projection model: \( f_{c,t+1} = f_{c,t} - d(\theta, f_{c,t}) \)

- \( f_{c,t} \) the TFR for country \( c \), 5-year period \( t \)
- \( d(\theta, f_{c,t}) \) the 5-year decline given by decline function \( d(\cdot, \cdot) \)

New extension: probabilistic projection model

1. Estimate \( \theta \) in \( d(\theta, f_{c,t}) \) for each country
2. Include uncertainty assessment:
   • Allow for random distortions
   • Assess uncertainty in $\theta_c$

Random walk with drift:

$$f_{c,t}^{t+1} = f_{c,t}^t - d(\theta_c, f_{c,t}^t) + \varepsilon_{c,t}$$

with
- $f_{c,t}^t$ TFR for country $c$, 5-year period $t$
- $d(\theta_c, f_{c,t}^t)$ 5-year decrement
- $\varepsilon_{c,t}$ Random distortions
Bayesian Hierarchical Model (BHM)

Bayesian inference: unknown parameters have probability distributions, which are “updated” with new information (prior distribution + data and model → posterior distribution)

Exchange information between countries using a hierarchical model:
- Unknown decline parameters are distributed around a “world average”
- For a specific country, its parameters estimates are determined by its observed declines, as well as the world level experience

Example: maximum 5-year decrement $d_c$
- Use a transformation of $d_c$ to restrict it to between 0.25 and 2.5 child:
  \[
  d^*_c = \log \left( \frac{d_c - 0.25}{2.5 - d_c} \right)
  \]
- Assume that $d^*_c$'s are exchangeable between countries
  \[
  d^*_c \sim N(\chi, \psi^2)
  \]
  with $\chi$ the world mean, and $\psi^2$ the variance of the $d^*_c$'s
The model is given by:

\[ f_{c,t+1} = f_{c,t} - d(\theta_c, f_{c,t}) + \varepsilon_{c,t} \]

Hierarchical distributions for country-specific parameters \( \theta_c \)

Prior distributions on the hierarchical parameters, and variance parameters of the distortion terms

Use Markov Chain Monte Carlo (MCMC) algorithm to get many samples of the set of model parameters

Each set of model parameters gives a future TFR trajectory

Many sets

→ Many TFR trajectories
→ Median projection and projection intervals
To get future $f_{c,t+1}$ for country $c$ in Phase II:
- Outcome of $\theta_c$ gives decrement $d(\theta_c, f_{c,t})$
- Sample a distortion $\varepsilon_{c,t}$ (use outcomes of its variance parameters)
- $f_{c,t+1} = f_{c,t} - d(\theta_c, f_{c,t}) + \varepsilon_{c,t}$

Repeat until start of Phase III:
earliest $t$ such that $\min\{f_{c,s} : s = 1, \ldots, t\} \leq \Delta_{c4}$ and $f_{c,t} > f_{c,t-1}$
Outline

1. Phase II: The fertility transition
2. Phase III: Post-transition low fertility
3. Results
Phase III: What happens post-fertility-transition?

Assume TFR will fluctuate around 2.1, use an AR(1) model:

\[ f_t = f_{t-1} + (1 - \rho)(2.1 - f_{t-1}) + e_t \]

\[ e_t \sim N(0, \sigma^2) \]

with \( \rho = \) autoregressive parameter with \( |\rho| < 1 \)
and \( \sigma = \) standard deviation of the random errors
**AR(1):** \[ f_t = f_{t-1} + (1 - \rho)(2.1 - f_{t-1}) + e_t \]

\[(f_t - 2.1) = \rho(f_{t-1} - 2.1)\]

**Fertility change for phase III**

- \( \mu = 2.1 \)
- \( \rho = 0.89 \)
- \( adj. R^2 = 0.95 \)

\[ f_t = \mu + \rho(f_{t-1} - \mu) \]
TF projection < 1.5 in 2005-2010

- Asymptotic 95% projection interval (PI) given by [1.6, 2.4] in 2100
- Use all below-replacement TFRs to estimate uncertainty in long-term projections ($s^2 = 0.27$ after 4 periods in phase III)
- Increased uncertainty in long range Phase III projections
No difference whether $\mu = 2.1$, 1.85 or 1.6
Outline

1. Phase II: The fertility transition
2. Phase III: Post-transition low fertility
3. Results
TF projection > 2.1 in 2005-2010
TF decline curve (country-specific)

Uganda

Guatemala

India

Venezuela (Bolivarian Republic of)
Out-of-sample model validation (2008 rev.)

Use data until 1980, and project until 2005-2010:

Summary of model validation results:

<table>
<thead>
<tr>
<th>Project</th>
<th>Above Median</th>
<th>Coverage 95%PI</th>
<th>Coverage 80%PI</th>
</tr>
</thead>
<tbody>
<tr>
<td>from 1980</td>
<td>43%</td>
<td>91%</td>
<td>77%</td>
</tr>
<tr>
<td>from 1995</td>
<td>36%</td>
<td>93%</td>
<td>79%</td>
</tr>
</tbody>
</table>
Out-of-sample model validation (2010 rev.)

Use data until 1990, and project until 2095-2100:

- **Uganda**: Over-predicted decline
- **Guatemala**: Under-predicted decline
- **India**: Under-predicted decline
- **Venezuela (Bolivarian Republic of)**: Predicted decline
Out-of-sample model validation (2010 rev.)

Use data until 1990, and project until 2095-2100:

- Under-predicted decline
- Under-predicted decline
- Under-predicted decline
- Under-predicted decline
Summary Bayesian TFR Projection Model

Probabilistic projection model for 5-year changes during and after the fertility transition

During the fertility transition:

- the 5-year decreases are modeled as a function of TFR level and decline parameters, with random distortions added to it
- the decline parameters are estimated with a Bayesian hierarchical model

After the fertility transition the TFR will converge to/fluctuate around 2.1, using an AR(1) model

Results: Country-specific projections that include an uncertainty assessment
References


White Paper: Probabilistic Projections of the Total Fertility Rate for All Countries for the 2010 World Population Prospects
Adrian E. Raftery, Leontine Alkema, Patrick Gerland, Samuel J. Clark, Francois Pelletier, Thomas Buettner, Gerhard Heilig, Nan Li, Hana Sevckova. (United Nations population Division, Expert Group Meeting on Recent and Future Trends in Fertility, New York, 2-4 December 2009)
