



# IMF Working Paper

---

## Avoid Filling Swiss Cheese with Whipped Cream: Imputation Techniques and Evaluation Procedures for Cross-Country Time Series

*Michaela Denk, Michael Weber*

## IMF Working Paper

STA

### Avoid Filling Swiss Cheese with Whipped Cream: Imputation Techniques and Evaluation Procedures for Cross-Country Time Series

Prepared by Michaela Denk, Michael Weber

Authorized for distribution by Ann McPhail

June 2011

#### Abstract

International organizations collect data from national authorities to create multivariate cross-sectional time series for their analyses. As data from countries with not yet well-established statistical systems may be incomplete, the bridging of data gaps is a crucial challenge. This paper investigates data structures and missing data patterns in the cross-sectional time series framework, reviews missing value imputation techniques used for micro data in official statistics, and discusses their applicability to cross-sectional time series. It presents statistical methods and quality indicators that enable the (comparative) evaluation of imputation processes and completed datasets.

JEL Classification Numbers: C13, C52, C59, C80

Keywords: Missing or incomplete data, imputation quality, statistical matching

Acknowledgements: Special thanks go to our colleagues at the IMF Statistics Department, especially Mike Seiferling, as well as at the Social Protection & Labor Unit of the World Bank's Human Development Network for valuable feedback and discussions.

Authors' E-Mail Addresses: [mdenk@imf.org](mailto:mdenk@imf.org), [mweber1@worldbank.org](mailto:mweber1@worldbank.org)

This Working Paper should not be reported as representing the views of the IMF. The views expressed in this Working Paper are those of the author(s) and do not necessarily represent those of the IMF or IMF policy. Working Papers describe research in progress by the author(s) and are published to elicit comments and to further debate.

Contents	Page
I. Introduction .....	<a href="#">3</a>
II. Data structures and missing data patterns .....	<a href="#">4</a>
III. Missing data techniques .....	<a href="#">6</a>
A. Traditional approaches .....	<a href="#">7</a>
B. Statistical Matching.....	<a href="#">10</a>
C. Multiple imputation.....	<a href="#">11</a>
IV. Applicability of missing data techniques to time series data.....	<a href="#">12</a>
V. Evaluation with statistical quality measures .....	<a href="#">15</a>
A. Degree of missingness .....	<a href="#">16</a>
B. Performance of imputation method.....	<a href="#">17</a>
C. Accuracy of imputation results .....	<a href="#">19</a>
D. Variability of statistics based on the imputed dataset .....	<a href="#">21</a>
VI. Conclusion .....	<a href="#">21</a>
VII. References .....	<a href="#">22</a>

## Figures

Figure 1. Missing data patterns for standard micro (=observation by variable) data .....	<a href="#">4</a>
Figure 2. Missing data patterns for multivariate time series and univariate cross-sectional time series .....	<a href="#">5</a>
Figure 3. Missing data patterns for multivariate cross-sectional time series .....	<a href="#">6</a>