Compliance and Usage in the Generations and Gender Programme

Arianna Caporali[1]

Institut national d'études démographiques, INED

Tom Emery

Netherlands Interdisciplinary Demographic Institute, NIDI

Abstract:

Launched in 2000 by the UNECE, the Generations and Gender Programme (GGP) is a longitudinal comparative survey of 18-79 years old in 19 countries in Europe and beyond run by a consortium of research institutions. It is based on a relatively decentralized management model and relies on considerable post hoc harmonization of data. The international "core" questionnaire is either adapted to the different national contexts or partly incorporated into existing surveys. Using data from the surveys administration, we examine the quality of compliance and standardisation in the GGP and whether this affects data usage. Firstly, we examine compliance by analysing the extent to which instruments from the core questionnaire were fielded within each of the 19 countries in the GGP. The results show that on average across countries, 66% of instruments in the core questionnaire were captured. Secondly, to examine usage, we take administrative data from the GGP website to capture the number of times each country dataset is downloaded. We supplement this with an analysis of the GGP bibliography and examine the number of times a country dataset is used in peer reviewed comparative publications (about 530 references). Finally, OLS regression analyses are presented to provide an overview of the association between compliance and usage, controlling for a number of contextual variables (e.g. number of IUSSP members, population with ISCED 8 in each country). The paper concludes with recommendations for future data collection activities and with reflections on the usefulness of analysis of compliance and usage in having an overview of the quality of comparative projects.

[1] Corresponding author : arianna.caporali@ined.fr

Introduction: the emergence of cross-country comparative surveys

In the last 30 years, there has been a growing number of "deliberately designed comparative surveys" (Harkness et al. 2010:4). Projects such as the International Social Surveys Programme (ISSP), the World Values Surveys (WVS), the Survey of Health, Ageing and Retirement in Europe (SHARE), the European Social Survey (ESS) have been run in an increasing number of countries. These projects are all based on a comparative methodology and allow for policy learning, economies of scale, and best practice sharing. Yet they may differ in the extent to which their design is deliberately comparative including such aspects as sampling, questionnaire design, data collection mode, documentation, or harmonization (Harkness et al. 2010). In general, such factors depend on the level of centralization of survey management. For example, surveys based on a centralized management model (e.g. SHARE and ESS) are more heavily involved in the implementation and enforcement of tendering and fieldwork guidelines to ensure that processes are comparable across countries. As a consequence, the comparability of a comparative survey requires its compliance levels to be demonstrated and this cannot be assumed simply on the basis of their methodological guidelines.

The case of the Generations and Gender Programme

Launched by the UNECE in 2000, the Generations and Gender Programme (GGP) is a longitudinal comparative survey of 18-79 years old in 19 countries in Europe and beyond run by a consortium of research institutions. It is aimed at studying the major demographic and economic challenges such as population ageing, low fertility, increased life expectancy and changes in family structures. To do so, it studies how micro-level factors influence the relationships between generations and gender along an individual's life-course and it complements the survey data with a contextual database (Vikat et al. 2007; Caporali et al. 2014).

The GGP is based on a relatively decentralized management model and relies on considerable post hoc harmonization of data. Defined by the central coordination team, the survey instruments and guidelines (UNECE 2005, 2007) are either adapted to the different national contexts or partly incorporated into existing surveys (such as in the case of Italy, Australia and Hungary). As of December 2015, data for the first waves for all countries have been made internationally available, as well as some wave two datasets. The GGP surveys have been used in over 500 peer reviewed publications.

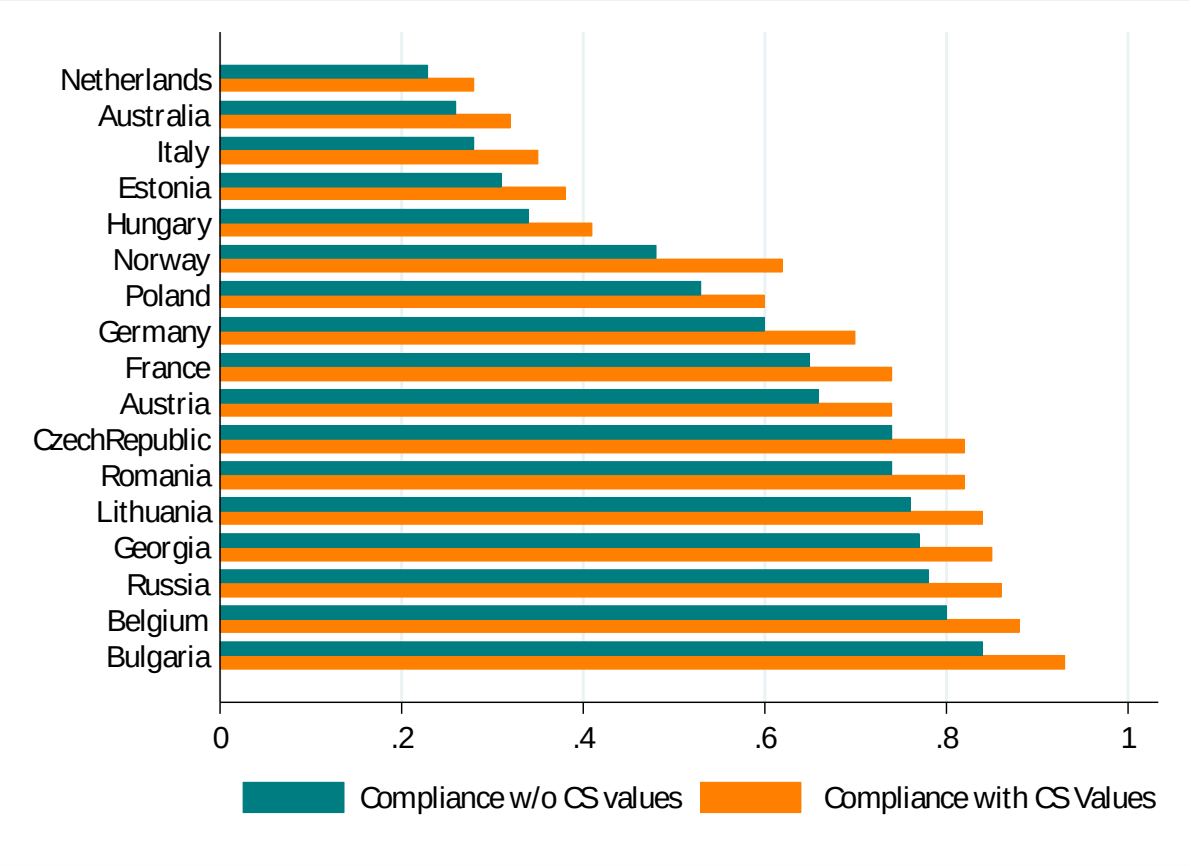
Previous works have focused on assessing GGP data quality, sampling and fieldwork methods (Demographic Research special issue, forthcoming, Fokkema et al. in particular). This paper examines the extent to which questionnaire standardisation affects data usage: Does compliance with the GGP "core" questionnaire lead to greater data usage and scientific output? We assume that compliance to the core questionnaire has no effect on the number of times a country-dataset is downloaded from the data access platform given that users would not know at the outset what the compliance level is. Instead downloads more accurately reflect demand for specific country data. Higher compliance levels, on the contrary, lead to more comparative publications.

Data and Method

We focus on the analysis of compliance and usage with reference to the first wave of the survey datasets available as of June 2015 (17 countries). We use three sets of data. First, based on the total number of variables in the pooled dataset, we calculate the compliance to the "core" questionnaire as the percentage of variables included in each country dataset. We distinguish between two types of compliance: a) compliance that excludes variables having country-specific response categories, b) compliance that includes variables having country-

specific response categories (Figure 1). In both types of compliance we do not consider the variables that were asked only in specific countries.

***Figure 1*: Overview of compliance (%) to the "core" questionnaire in GGP wave 1 datasets**



The second set of data concerns data usage. We consider the number of times each country dataset has been downloaded from the GGP data platform. We also examine the number of peer reviewed publications (i.e. journal articles and book chapters) recorded via the GGP website. We code the GGP bibliography to count the number of times a country dataset is used in comparative publications. We plan to supplement this with the total GGP citations via Google Scholar.

The third set of variables is derived from GGP data documentation and from various international sources (e.g. Eurostat, the World Bank, UNESCO). This information is used as controlling variables in our analysis. We consider: the year of fieldwork and the number of months the country-datasets have been available, the number of IUSSP (International Union for the Scientific Study of Population) members per country, population with ISCED 8, total population, and GDP per capita.

We run linear regressions to assess the effect of compliance on downloads and on the number of comparative publications. We also consider each section of the questionnaire separately and we estimate their contribution to comparative publications.

Analysis

On average across countries, 66% of instruments in the core questionnaire were captured. This ranges from 23% (the Netherlands) to 84% (Bulgaria). The countries with the highest compliance are not necessary those with the greatest number of downloads and comparative publications. For example, Germany and France which rank in the middle as to

compliance to the "core" questionnaire (respectively 60% and 65%), yet are among the countries with the greatest number of downloads and publications.

Our regression models reveal that the number of months a dataset has been available has the strongest significant effect on both downloads and number of comparative publications. As expected, compliance does not affect downloads while it as a positive and significant effect on the number of comparative publications. The inclusion of the other variables in the models does not change the regression coefficients of compliance. The analysis by section of the questionnaire reveals that the sections about fertility, health and income appear to have a strong association with scientific output.

Conclusion

Initial analyses show that compliance does not affect downloads, while for every 2% of compliance there is an additional publication. Compliance to the questionnaire together with rapid availability of the datasets to the research community appears to be key component of success in terms of scientific output.

The analysis of compliance and usage is a key element to assess the success of deliberately designed comparative surveys. This is especially true for international survey programmes based on relatively decentralized models.

References

Caporali, A., Klüsener, S., Neyer, G., Krapf, S., Grigorieva, O. (2014). The contextual database of the Generations and Gender Programme: Concept, content and research examples. Stockholm University, Stockholm Research Report in Demography, 6.

Harkness, J.A., Braun, M., Edwards, B., Johnson, T.P., Lyberg, L.E., Mohler, P.Ph., Pennell, B., Smith, T.W. (2010). Comparative Survey Methodology. In: Harkness, J.A., Braun, M., Edwards, B., Johnson, T.P., Lyberg, L.E., Mohler, P.Ph., Pennell, B., Smith, T.W. (ed.). *Survey Methods in Multicultural, Multinational, and Multiregional Contexts*, Hoboken, New Jersey: Jhon Wiley & Son: 3-16.

Vikat, A., Spéder, Z., Beets, G., Billari, F., Bühler, C., Desesquelles, A., Fokkema, T., Hoem, J.M., MacDonald, A., Neyer, G., Pailhé, A., Pinnelli, A., and Solaz, A. (2007). Generations and Gender Survey (GGS): Towards a better understanding of relationships and processes in the life course. *Demographic Research* 17(14): 389–440.

UNECE – United Nations Economic Commission for Europe. (2005). *Generations and Gender Programme - Survey Instruments*. New York/ Geneva: United Nations.

UNECE – United Nations Economic Commission for Europe. (2007). *Generations and Gender Programme - Concepts and Guidelines*. New York/ Geneva: United Nations.

Fokkema, F., Kveder, A., Hiekel, N., Liefbroer, A., and Emery, T. (Forthcoming). The Generations and Gender Programme Wave 1: Sampling and Weighting. *Demographic Research*.