

## RESPONDING TO NEW DATA DEMANDS FOR COMPARATIVE RESEARCH AND MULTILEVEL ANALYSIS: THE CONTEXTUAL DATABASE OF THE GENERATIONS AND GENDER PROGRAM

Martin Spielauer<sup>1</sup>

### ABSTRACT

The increasing recognition that the study of human behaviors has to take into account the multiple contexts in which they occur has opened a promising research avenue in social sciences. It also presents new challenges, i.e. to complement micro-level surveys with the collection of meaningful contextual data within a common conceptual framework. The international comparative Generations and Gender Program, which is coordinated by the Population Activities Unit of the United Nations' Economic Commission for Europe, combines a panel survey carried out in various European countries, Japan, and Australia with a comparative contextual database developed as integral part of the program.

**KEY WORDS:** Database; Multilevel; Comparative; Generations; Gender

### 1. INTRODUCTION

This paper describes the conceptual framework and information content of the international Contextual Database of the Generations and Gender Program (GGP). The Contextual Database was developed as complement to the core element of the program, the Generations and Gender Survey (GGS), a panel survey with panel intervals of three years. The GGP is expected to generate the main data source for the cross-national, comparative, and multidisciplinary study of the dynamics of family relationships in a series of Eastern- and Western European Countries, Japan, and Australia. Currently, data of the first wave are becoming available for the first participating countries, others are to follow in the next few years. To date (2006), the Contextual Database provides data for Bulgaria, Russia, and Hungary; several other countries are scheduled to follow in 2006 and 2007.

The contextual database aims at responding to the data demands for comparative research and multilevel analysis by providing a comparative collection of around 210 variables on a national and sub-national level, complementing the individual-level data collected in the GGS. It includes information on laws and policies that effect age and sequencing norms as well as the consequences and risks associated with the central life course events studied in the GGP. It also contains indicators of general gender and generational relations, and a range of social, economic, and political indicators. The contextual database allows reaching beyond the mere analysis of single countries by considering the extent to which differences in contexts shape the demographic processes, and offers the opportunity to conduct comparative studies on the interaction between the micro and macro dimension.

This contribution is organized as follows. First, we give a structured description of the GGS information content to allow the matching of micro-level survey information with corresponding macro-level contextual domains that constitute the topics of the Contextual Database. Second, we highlight how the Contextual Database meets some of the key demands of the contextual and multilevel analysis by allowing to link individuals to context over time, by region, or by population group. Finally, we describe the key measurement dimensions of the Contextual Database. These three parts correspond to three of the main development goals of the Contextual Database. These are: deep integration into the Generations and Gender Program, the support of innovative methods of data analyses, and a theory-based conceptual framework. The Contextual Database is freely available at the website of the Max Planck Institute for Demographic Research (Note 1).

---

<sup>1</sup> Max Planck Institute for Demographic Research, Konrad-Zuse Strasse 1, 18057 Rostock, Germany.  
[spielauer@demogr.mpg.de](mailto:spielauer@demogr.mpg.de)

## 2. MICRO-MACRO MAPPING AND DATABASE TOPICS

In this chapter, we present the 16 topics or domains of the Contextual Database and their relation to the GGS. The selection process of appropriate database topics is based on a structured inventory of the micro level information collected by the survey in a way that supports the identification of higher-level “systems” corresponding to micro-level domains and relations. One way that has proven fruitful in this respect is to view the survey content from a life course perspective. This provides a very systematic approach regarding the structuring of complex individual data by different life course domains (roles) and by time. Three survey focal points that are not fully captured by the life course approach and therefore studied separately are intergenerational and gender relations, income and wealth, and values.

The GGS is the successor of the Family and Fertility Survey (FFS) which was carried out in around 20 mostly European countries in the early 1990s. The GGS is a new round of comparative surveys to gain a deeper understanding of the inter-generational and gender relationships as well as the relationship between social, economic, political, and ideational change and demographic behavior. Like the FFS, GGS has a strong retrospective component – it collects event history data on fertility, partnerships, education, and economic activity. In difference to the FFS, it has a prospective component, first by including intentions concerning most studied events, and second by its design as a panel survey. Another difference to the FFS is the full age range of the GGS: The survey includes men and women aged 18-79. Sample sizes vary by country and range between 10.000-16.000 individual respondents. The Generations and Gender Program is developed by a consortium of leading Demographic Institutes (Note 2). By the end of 2005, around 15 Western and Eastern European countries, Japan, and Australia indicated their participation (Note 3). The main dependent variables of the survey are childbearing, partnership formation and dissolution, transition to adulthood, living arrangements, and retirement. The broad range of covariates includes: activity, education, income and transfers, assets, value orientations and attitudes, intentions, contraception, household organization, social networks, childhood context, health & well-being and – as indicated by the title of the program: generations and gender aspects.

**The individual life course and its context:** The life course refers to ‘a sequence of socially defined events and roles that the individual enacts over time’ (Giele and Elder 1998, p. 22). These life course events are central issues in the GGS and constitute key dependent variables. The life course perspective also goes hand in hand with a certain description of human biographies as event histories, with clearly defined states and events belonging to different domains. The description of individual lives by means of event history data is not independent of the context. For example, the activity status “on parental leave” depends on the existence of parental leave regulations. This places individual choices into a context that may vary considerably between countries. In the GGS, we can distinguish several such domains, with most of their various states and dimensions displaying a direct link to contextual domains.

A first group of life course information collected in the GGS concerns demographic events. Here, births and partnership transitions constitute central dependent variables; the survey contains full birth and partnership histories. Fertility and partnership events are related to many contextual domains of the database, and a hypothesis linking contextual variables to these events constitutes one of the main selection criteria of variables. The health dimension is another domain closely related to life itself. In the GGS, data is collected on current illness, disabilities (including some retrospective information), care needs, and wellbeing. Data on limitations in daily activities are collected for all household members as well as for parents and children in order to obtain information on possible care needs in the family network. Corresponding contextual information can be found under the headings:

- **Topic 1 - General demographic indicators:** a brief selection of key demographic indicators on the national and regional level. The behavioral focus lies on the classic demographic events of birth, marriage, divorce, migration, and mortality; additional topics are abortion and single motherhood.
- **Topic 2 - Health and health care system:** the contextual variables include measures on healthy life expectancy as well as general, maternal, and infant mortality. We also aim at describing the main features of the national health care system.

An important life course domain in the GGS concerns activities. The survey distinguishes between various main activities, such as working, retirement, on parental leave, unemployed, in education, in national service. Information on the main activity is collected for all household members. Detailed questions regard the respondent's and partner's current job and/or education and intentions to start or resume work. Full activity histories are to be collected during the second wave. Again, we find a close link between individual states and contextual domains, which are:

- **Topic 3 – Labor and employment:** a selection of indicators including gendered labor-market participation, sectoral employment, wages and wage distributions, work time regulations, policy information on the promotion of part-time work, legal entitlement to part-time work, standard hours, extra compensation for work during non-standard hours, and measures to reduce work during non-standard hours.
- **Topic 4 – Unemployment and unemployment system:** we focus on three aspects of unemployment: unemployment rates by individual characteristics such as age, sex, and education; measures on the duration of unemployment; and the main characteristics of the unemployment insurance system.
- **Topic 5 – Pension system:** indicators focus on retirement age and system descriptions, e.g. the link of contributions to benefits and the extent to which child and family care are accounted for.
- **Topic 6 – Parental and care leave systems:** the database provides a standardized table on all national maternity, parental, and childcare leave systems. Indicators include eligibility, timing, cash compensation, and the flexibility of leave systems.
- **Topic 7 – Military system and national service:** the indicators focus on the service age and service durations of compulsory military and civilian services and the range of choices concerning service type and timing.
- **Topic 8 – Education:** the data cover various aspects of education, including a description of the national school system, timing norms like entry age and school durations, standard school hours, enrollment rates and educational attainments, school finance, and school quality.

The **socio-economic situation** of the respondent and the household is usually closely linked to employment and other activities. The GGS collects information on labor and transfer incomes, wealth, and some selected expenditures on care and household services. This information is complemented by contextual data of two domains:

- **Topic 9 – Economy:** this database section consists of a selection of general economic indicators that are not directly linked to specific activity statuses: economic output, growth, stability, income distribution, and poverty.
- **Topic 10 – Housing:** Information on housing stock and housing production by type, general housing situation, and relevant housing policies

With it being a “Generations and Gender” Survey, **gender and (inter-)generation relations** constitute a key area of information. Data on relations are collected for persons in the household, all (ex-) partners, children and parents as well as other persons if they are important providers or receivers of care. Personal relations are investigated in various dimensions, including their legal nature, co-residence, intensity, satisfaction, quality, power of decision making, money, and time. One of the key areas of relations is the organization, interchange, and provision of care, which also constitutes a central link between gender and generational relations. The Contextual Database complements the individual-level information in the following fields:

- **Topic 11 – Legal relations of personal relations:** including information on marriage and divorce regulations and restrictions, restrictions of abortions and legal obligations concerning care.
- **Topic 12 – Childcare:** contextual data concern the availability of care institutions as well as costs, and the extent to which they are shared between parents and the public, the quality of childcare institutions, the existence of legal entitlements to institutional care, enrollment rates, and timing norms such as (pre)school entry ages and typical school hours.

- **Topic 13 – Elderly care:** this database section contains measures that support elderly in need for care, supportive measures for care givers, and general indicators on people receiving institutional care of different types.
- **Topic 14 – Tax-Benefit system:** including tax rates, child benefits, and the impact of marriage on taxes and benefits.

The GGS includes a section on attitudes and **value orientations** on gender and intergenerational relations as well as on religiousness. This individual level data is placed in the context of the cultural, political, and religious heritage of a country or region:

- **Topic 15 – Culture:** selected indicators on the religious, linguistic, and ethnic composition of the population.
- **Topic 16 - Political system:** information on the main political parties and government coalitions.

All variables belong to one of the 16 contextual domains and focus on one or more of the key dependent behaviors (fertility, partnership, transition to adulthood, living arrangements, and retirement), with special emphasis on generational and gender aspects.

### 3. THE MICRO-MACRO LINK IN DATA ANALYSIS

With a view to support a multilevel approach to GGP data, we have to ensure that it is possible to link individual level data to their relevant context. Such a link can be established by time, region, or individual characteristics such as sex, education, or eligibility to policies. In order to meet the historical depth of the GGS, the Contextual Database provides around 65 national level **time series** and 70 **policy histories** ranging back to 1970. Time series are usually yearly numerical indicators; policy histories give text and standardized numeric descriptions of key policy changes and changes in legal regulations.

A link by region can be established for cross-sectional data and concerning the prospective component of the survey. Around 60 variables of the contextual database are collected on the **regional level** and will be updated with the successive waves of the survey. In order to capture recent regional trends, many of the regional level variables are collected from 2000 onwards. The choice for an appropriate regional level is taken individually by country, taking into account sample size requirements of multi-level models (as treated in Racioppi and Rivellini, 2001), the availability of regional data according to national statistical systems, and the regional identifier available in the national GGS.

Besides the general link of individuals to context by time and region, different population groups find themselves in very specific contexts. Typical population groups frequently separated in aggregated statistics are constructed by gender, age group, marital status, educational attainment, and occupation. In order to determine the policy contexts for individuals of different characteristics, the description of policy measures emphasizes especially the concepts of eligibility.

The CDB is a **comparative database**: the same set of variables is collected for all participating countries, following the same definitions. In practice, this has proven difficult for various reasons. For many variables, there exist competing definitions used in different international databases covering different sets of countries. For other variables, international definitions became standards only recently, not allowing the collection of time-series data according to them. As a general rule, we aim at capturing variations over time as fully as possible; this, however, sometimes restricts us to use the national definitions for which such data are available only. Shorter, comparable time series are then provided additionally.

#### 4. THEORETICAL CONCEPTS OF THE MICRO-MACRO LINK

In the Contextual Database, we distinguish between two partly overlapping theoretical concepts of context assumed to influence individual behavior: the normative one – statistical and legal norms – and the welfare state context, expressed in policies that affect risks, agency, equality, and social rights.

Statistical norms refer to regular behavior; a behavioral pattern becomes a norm if the majority of actors behave according to this pattern. Thus, norms are objectively observable and measurable, i.e. derived by aggregation describing means and distributions of individual events and characteristics. In difference to this descriptive concept of norms, legal norms are prescriptive and value-oriented: they are social facts that can be identified through the existence of certain sanctions rather than directly.

In the contextual database, we include **statistical norms** in the form of mean values and age distributions of the key demographic events, such as the mean age at childbearing and marriage, assuming that individuals orient their behavior on such norms. In addition, many economic indicators fall into this group. As statistical norms change over time, they can also serve as indicators of the developmental stages, assuming universal transition mechanisms as in demographic transition theory. In this sense, statistical norms can be used as indicators of the prevalence or emergence of certain behaviors (e.g., unmarried cohabitation). Statistical norms are also present in the context of policies, where they constitute important outcome variables, e.g., measuring enrolment and take-up rates.

**Legal regulations** influence individual behaviors in various ways. They impose important timing norms that impact key domains of the individual life course. This is especially visible for education systems that e.g., define the minimum compulsory time of schooling. Important timing norms are also set by work time regulations, paid maternity leave periods, the conscription age, and time of obligatory military services or the legal retirement age. Legal norms also impose important quality standards, e.g., staff requirements of childcare institutions, which may impact individual decisions on the organization of care. They additionally define family responsibilities and may restrict, prohibit or regulate certain behaviors and personal relations with examples being abortion laws, divorce regulations, and the legal treatment of same-sex partnerships.

The **welfare state context** is expressed in policies and their underlying social values and concepts. Besides setting norms, the welfare state influences individual behavior through four main channels: it defines social and economic rights; it influences (equality-) inequality structures; it alters the risks and consequences of certain events; and it influences the individual agency. (A detailed description of welfare state concepts relevant to the database is given in Neyer (2003)).

The term “welfare state” is closely linked to Marshall’s (1950) concept of social citizenship based on the recognition of material and social needs to give rise to economic and social **rights**. An influential way of specifying social citizen rights is linked to Esping-Andersen’s concept of de-commodification: the weakening of the connection between income and market participation (Esping-Andersen 1990). The degree to which policies and social security systems are designed in recognition of social and economic rights is reflected in legal entitlements, eligibility, coverage, the linkage of benefits to individual contributions, and the existence and level of minimum standards, e.g., minimum benefits or the quality standards of services.

The concept of **equality** refers to the cross-sectional differentiation of living conditions and the equality of life chances, i.e., the distribution of opportunities, resources and capabilities, including their intergenerational transmission. Equality has important gender, generational, and class dimensions. Measures of the extent of equality are participation and enrolment rates (e.g., labor-market participation by sex and age) and distributions of e.g., income. Policy dimensions reflecting equality concepts include benefit eligibility and coverage, public sector involvement in service provision (e.g., schools), the existence and mechanisms of measures to provide access to services and markets (e.g., housing, education), the treatment of unmarried and same-sex partnerships, the special protection of institutions, such as the family, and affirmative actions to actively promote equality.

**Risks** and security have to do with the likeliness of unfavorable events, such as unemployment and sickness, the consequences of such events, such as poverty and dependence, how these risks and consequences are related to the life course events and relations studied in GGP, and the ways in which the state impacts these risks and consequences. Risks directly addressed in the database are poverty and unemployment. Important policy dimensions

are replacement rates including the underlying concept (e.g., poverty prevention vs. status maintenance), durations (of benefits; average durations of unemployment) and the treatment of unpaid care work by the pension system.

**Agency** refers to a set of choices, the ability to choose and to the incentive system that influences the choice to be made, including the individual's bargaining position within the family. The concept of agency is frequently combined with the concept of equality in terms of 'equality in the freedom to achieve' and is especially useful for the study of gender differentials (Korpi 2000). Policies interfere with agency in various ways that are addressed in the Contextual Database: the economic burden and time commitment associated with behaviors studied in the GGP: e.g. costs and opening hours of childcare institutions; the flexibility of policies (influencing the set of choices, e.g., the deferability of parts of leave periods); and incentives to promote gender equality, e.g. in the organization of paid and unpaid work and the take-up of leave periods.

## 5. SUMMARY

The Generations and Gender Contextual Database is a theory-based instrument of the Generations and Gender Program complementing the individual-level data collection of the Generations and Gender Survey. The database provides time series, policy history and regional level data on sixteen topics and is designed to support the integration of micro- and macro level approaches of analysis. It is freely available at the website of the Max Planck Institute for Demographic Research. (Note 1)

## NOTES

- Note 1: The Contextual Database and additional information on the program can be found at the website of the Max Planck Institute for Demographic research: <http://www.demogr.mpg.de>
- Note 2: For detailed information, consult the GGP website of the Population Activities Unit of the United Nations' Economic Commission for Europe: <http://www.unece.org/ead/pau/ggp/Welcome.html>
- Note 3: Participants of the GGP by the end of 2005: Belgium, Bulgaria, Czech Republic, Estonia, France, Georgia, Germany, Hungary, Japan, Norway, Poland, Romania, and Russia. Additionally, Australia and Italy have incorporated parts of the questionnaire into existing surveys.

## REFERENCES

- Esping-Andersen, Gøsta (1990) *Three worlds of welfare capitalism*; Polity Press, Cambridge
- Giele, Janet Z. & Elder, Glen H. (Eds.) (1998) *Methods of Life Course Research: Qualitative and Quantitative Approaches*, Thousand Oaks, Sage
- Korpi, Walter (2000) *Faces of inequality: gender, class, and patterns of inequalities in different types of welfare states*; Social Politics, Summer 2000
- Marshall, Thomas Humphrey (1950) *Citizenship and social class and other essays*. Cambridge, Cambridge University Press
- Neyer, Gerda (2003) *Gender and generations dimensions in welfare-state policies*, MPIDR Working Paper WP-2003-022. <http://www.demogr.mpg.de/papers/working/wp-2003-022.pdf>
- Racioppi, Filomena & Rivellini, Giulia (2002) *The contextual dimension in GGP: some methodological issues about data collection and sampling procedures* <http://www.unece.org/ead/pau/ggp/smiwg/cdbasegr.pdf>