Research Project AG/01/070

Validation of the individual part of the social security concept matrix

Summary

prof. dr. J. Berghman & Kristine Nijs

Social Policy Unit
CESO (Centre for Sociological Research)
K.U.Leuven

Commissioned by

Contents

1. The goal of the research project

The Agora research project AG/01/070 'Validation of the individual part of the social security concept matrix' (the Validation Project) is a direct continuation supplementary to the Agora research project AG/01/039 'Harmonisation of statistical concepts in the social security' (the Harmonisation Project). Both research projects (for the period September 1999 - December 2001 and the period January 2002 - December 2004) were conducted by the Centre for Sociological Research (Social Policy Unit, K.U.Leuven) under the direction of Prof. Berghman and were commissioned by the Federal Public Planning Service (FPPS) Belgian Science Policy and the Federal Public Service (FPS) Social Security.[1]

The main aim of these research projects is to develop *a consolidated overall picture of the Belgian population in terms of social security status* (Berghman & Nijs, 2003; Berghman, Nijs & Nachtergaele, 2004). In Belgium, several social security institutions are involved in implementing social policy. To mention are institutions that collect social contributions, the diverse institutions that grant benefits to the beneficiaries and the institutions responsible for inspection on all of that. All those institutions each have their own databases with administrative data they need to perform their specific tasks. Hence there is an extremely amount of information on social security in Belgium, but this information is spread over various databases with each their own particularities. The aim of our research project is to develop an instrument that can gather all relevant social security data in a harmonised and structured way so that they can be managed consistently and usefully. This should allow the FPS Social Security to gain accurate insight on social security in Belgium by means of better statistics and innovating analyses and publications. They are in need of a good instrument in order to be able to respond more easily to various questions presented to them regarding all kinds of themes in the field of Belgian social protection. More precisely, they demanded us for an instrument that can reveal detailed insight on beneficiaries and contributors. Moreover, in order to evaluate past and current policy and in order to develop proper future policy, the FPS Social Security needs other kinds of information, supplementary to the existing available information. Instead of overall expenditure and receipt figures for the diverse branches of social security, one must have information on the contents and characteristics of

---

[1] Research reports on these projects (Berghman, Booghmans, Boos, Nachtergaele, Verhalle & Vleminckx, 2004; Berghman & Nijs, 2004; Berghman & Nijs, 2005) are available at the Centre for Sociological Research (CESO) at the K.U.Leuven.

those social policy measures as well. For example what part of total expenditure is wage related? Furthermore it would be interesting not only to present analyses on the amount of benefits by gender or age, but also to present information on the (past) labour market situation or the household setting of beneficiaries. A final aspect we explicitly take into account is that the requirement of the new instrument to comply with the relevant methodologies and data demands of international organisations like Eurostat, OECD, United Nations and ILO. This must allow the FPS Social Security to provide correctly the demanded data to those organisations.

In order to achieve those objectives the research team has developed an overall framework to describe and to monitor the whole social security system from all angles in a structured and consistent way. This conceptual outline we call the social security concept matrix (or *social security matrix* in short). The framework of the matrix shows in which way relevant information on Belgian social protection (characteristics of social security schemes and their expenditures and receipts) and on Belgian population in terms of social protection status (protected persons, beneficiaries and contributors) can be captured and managed in a database. The social security matrix consists of three parts. In the *product matrix*, social security products[2] are coded on variables that characterise those products and also expenditure and receipt figures are taken in. In the *individual matrix*, persons are characterised by variables not only relating to individual properties but also to their household setting. Finally, the *connecting matrix* relates individuals to the social security products they are protected by, benefit from and/or contribute to.

Within the Harmonisation Project the product matrix was developed. In addition to that, the main aim of the Validation Project was to further develop and to validate the individual matrix and the connecting matrix.

---

[2] In order to relate individuals to the social security measures that apply to them and in order to score social security measures on the variables of the product matrix, social security measures have to be broken down (disaggregated) to the level of the individual field of application. Therefore the unit of observation in the product matrix are not social protection measures but what we call social security products. A social security product is each social protection measure as well as each part of it disaggregated down to the individual level (Berghman, Booghmans, Boos, Nachtergaele, Verhalle & Vleminckx, 2004, pp. 59-60; Berghman & Nijs, 2005, p. 6).

2. Theoretical elaboration

The focal purpose of the research project is to reveal a consolidated picture of the Belgian population in terms of their social security status. To this end, in the first phase of the Harmonisation Project a new broad concept of social security was developed. In addition, explicit attention was paid to the compatibility requirements of international organisations such as Eurostat, OECD, United Nations and ILO regarding social statistics. We also draw some inspiration from other existing national models for data-management of administrative data in the realm of social protection.

In the second phase of the Harmonisation Project, the product matrix was constituted. Departing from theoretical insights, considering relevant aspects for policy research and taking into account international data demands on social statistics, the variables to be included in the product matrix were selected. The product matrix was also put to a first validation test by means of the social security products in the field of inability to work. Throughout the Validation Project, the product matrix was modified to some extent, guided by new developments of international reporting methodologies and by new insights from other research projects of our research team, especially regarding supplementary social provisions. Moreover, the product matrix was extended with financing measures and updated for the four quarters of 1999 to 2004 (cf. paragraph 3.1).

In addition to that, also the individual matrix and the connecting matrix were developed within the Validation Project. By means of the connection of administrative data from various social security institutions realised by the Crossroads Bank for Social Security, the individual matrix and the connecting matrix were implemented as much as possible in the form of a database on the personal level (cf. paragraph 3.2).

2.1. A broad concept of social security

The research team defines social security as *all redistributive measures, aimed at income protection and health care, which are collectively recognised* (Berghman & Nijs, 2005, p. 3). It encompasses protection against social risks on which factions (groups) in society agreed upon that responsibility for those risks and their consequences should be shouldered totally or partially by society as a whole. In operationalising this concept of social security, we consider a broad set of policy instruments. Since for an accurate and complete picture of social security, the following three financial flows have to be envisaged (Adema, 2001): the social benefits to individuals and households, the means to finance these benefits and the tax breaks with a social purpose. Moreover, not only public provisions but also occupational and private provisions and the impact of the tax system[3] must be taken into account as well. Hence, the social security matrix is not restricted to first pillar provisions. Also second and third pillar provisions and fiscal measures with a social purpose must be incorporated in the matrix.

---

[3] The tax system can affect social spending in three ways: (1) governments levy direct taxes and social security on cash transfers, (2) governments levy indirect taxes on goods and services bought by benefit recipients and (3) governments make use of the tax system to directly pursue social policy goals (Adema, 2001, pp. 14-15). This last category (tax breaks for social purposes) covers tax advantages (reductions, exemptions, deductions and postponements) similar to cash benefits and/or aiming to stimulate the provision of private social benefits (Adema, 2001, p. 21). Adema shows that for an accurate insight on social security the impact of the tax system has to be taken into account in order to move from gross to net (after tax) social expenditure statistics. In doing so, he operationalised some ideas that were presented in Berghman's oration (Tilburg, 1986).

2.2. International data demands on social statistics

To allow the FPS Social Security to correctly report the Belgian social security data to international organisations like Eurostat, OECD, United Nations and ILO, the social security matrix in general and the product matrix more specifically are constructed in such a way that the new instrument is in compliance with the reporting requirements put forward by these international organisations (Berghman, Booghmans, Boos, Nachtergaele, Verhalle & Vleminckx, 2004, pp. 21-43). First of all some important developments in the approach on social statistics were described. Within Eurostat two themes are of relevance in this context: the striving for harmonised and integrated (compatible) social statistics and the increasing use of administrative data and registers for statistical purposes. The European requirements are presented more concretely in the Esspros Manual (Eurostat, 1996). Esspros stands for 'European System of Integrated Social Protection Statistics'. Besides this European standard, the specific demands of OECD regarding social expenditure statistics were reviewed as well. The Social Expenditure Database 1980-1996 (SOCX Database) is a database containing data on social protection expenditure in the different OECD countries (OECD, 2001). An implicit complement to the SOCX Database is put forward in two recent publications of Willem Adema. The first investigates the impact of private social expenditures (Adema & Einerhand, 1998) and the second demonstrates the need for net instead of gross expenditure statistics (Adema, 2001). In addition, the reporting methodology of the International Labour Organisation was taken into account. The ILO surveys on a regular basis the cost (receipts and expenditures) of social security in its member states. In the inquiry into 'The Cost of Social Security' the ILO compares the gross direct social security receipts and expenditures of the different ILO countries over a period of one year (ILO, 1994). After an adjustment of the methodology in 1997, the inquiry ended in 1999. Currently the ILO is working on a new 'Social Security Inquiry' (SSI) in which besides social security expenditures and receipts data shall be gathered on the amount of beneficiaries, protected persons, contributors and average benefit amounts (ILO, 2004). Finally the system of national accounts was considered, more exactly the by Eurostat in 1995 revised European System of National and Regional Accounts (Eurostat, 1995).

Also for selecting the variables, concepts and classifications to be included in the individual matrix data demands and methodologies of international organisations were examined (Berghman & Nijs, 2004, pp. 30-56). Compliance with Eurostat requirements deserves prior attention. Harmonising social statistics is an important theme within Eurostat. This resulted among other things in the publication of several harmonised core variables on persons and households, with a description of the recommended concepts, definitions and measuring rules (Eurostat, 2000). Another inspiration while selecting variables in the individual matrix are Eurostat's recommendations regarding population censuses. In 1998 Eurostat and the United Nations Economic Commission for Europe jointly presented a series of recommendations for the population censuses in the ECE Region for the year 2000. These recommendations cover firstly a list of core topics that countries should include in their census as well as the recommended definitions and classifications for each of those core topics and secondly a list of non-core topics that countries may want to include in their census as well as the recommended definitions and/or classifications for some of these non-core topics (United Nations, 1998). One year later Eurostat (1999) published similar recommendations in the realm of the 'Community Programme of Population and Housing Censuses in 2001'. The joint recommendations of United Nations and Eurostat were its groundwork. The majority of definitions and classifications of the topics are the same as in the 1998 joint recommendations. Finally, also the new Eurostat project concerning 'Community Statistics on Income and Living Conditions' (EU-SILC in short) was considered in developing the individual matrix (Eurostat, 2001). Moreover, with respect to certain variables also other organisations are of great importance: for employment variables the International Labour Organisation is the utmost reference (ILO, 29.10.1982; ILO, 28.01.1993), for education variables is referred to the United Nations Educational, Scientific and Cultural Organization (UNESCO, 1997) and for income variables the recommendations of the Expert Group on Household Income Statistics (2001) were incorporated.

2.3. The framework of the social security matrix

In a next stage of the research project, the social security matrix was conceptually designed as an 'entity-relationship-model' (figure 1). Such a model points out how the connection between the product matrix and the individual matrix can be made and shows schematically how the social security matrix can be implemented in the form of a comprising database. In the social security matrix, various administrative data have to be captured in one data file. The data are to be stocked at a quarterly basis. As pointed out earlier, the social security matrix consists of three parts: the product matrix, the individual matrix and the connecting matrix. These characterise respectively social security measures, the Belgian population and the relation between both. Each of the separate matrices encompasses a series of variables that are divided into some modules as regards content.

In the *product matrix* (Berghman & Nijs, 2005, pp. 6-62) the social security product is the unit of observation. A social security product is each social protection measure as well as each part of it disaggregated down to the level of individual application modalities and benefit entitlements, conditions and provisions. Once all social security products are inventoried, they are scored on the variables of the basic module, the expenditure module and the financing module. The basic module comprises some general variables that are of relevance for each social security product and that stipulate whether the social security product further has to be coded either in the expenditure module or in the financing module. The expenditure module and the financing module then cover more specific variables concerning respectively expenditure measures and financing measures as well as the relevant amounts. The expenditure and financing figures in the product matrix have to refer to the total amount that was spend or received in the course of the quarter.

Figure 1: The social security matrix as entity-relationship-model
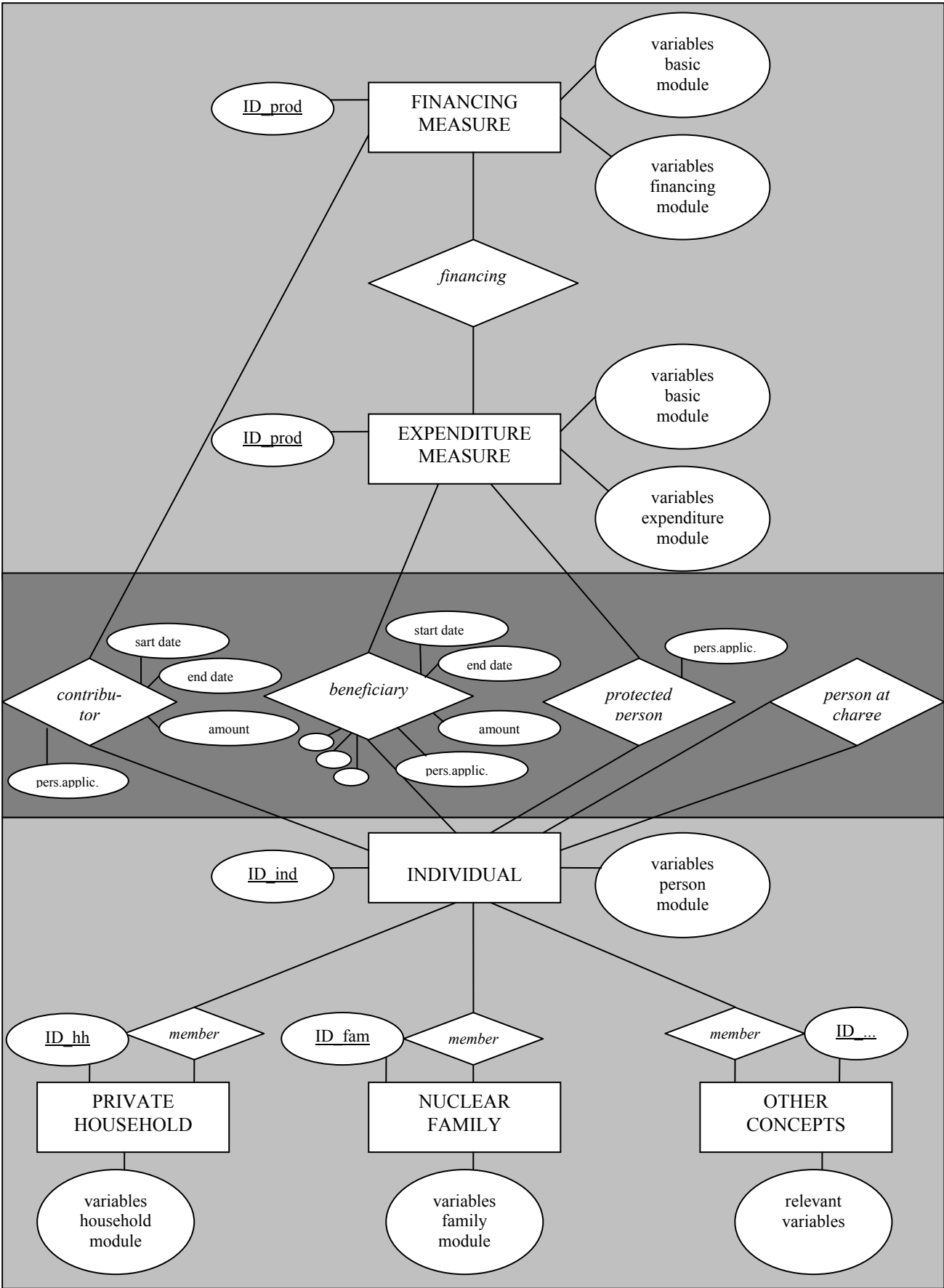
The variables in the *basic module* are 'type of social security product' (social expenditure measure, social financing measure, fiscal-social expenditure measure, social contributions on benefits, direct taxation on benefits, indirect taxation on benefits, indirect taxation on consumption), 'legal enforcement' (compulsory, compulsory but with possibility to opt-out, voluntary), 'level of decision-making' (European or other international government, federal government, state government, local government, social security institutions, sickness funds, public centres for social wellbeing, insurance and financial institutions, binding intersectoral CLA[4], non-binding intersectoral CLA, social partners intersectorally not by CLA, binding sectoral CLA, non-binding sectoral CLA, social partners sectorally not by CLA, CLA at corporation level, social partners at corporation level not by CLA, employer discretionary, other), 'level of financial administration' (European or other international government, federal government, state government, local government, social security institutions, sickness funds, public centres for social wellbeing, insurance and financial institutions, employers by internal financing, other), 'scope' (special scheme, general scheme, universal scheme, residual scheme), 'periodicity' (yearly, quarterly, monthly, weekly, daily, other periodicity, on a single specific occasion, non-recurrent, no determined periodicity), 'family related' (yes or no), 'means-tested' (yes or no) and 'evolution social security product' (not changed, changed, new, abolished).

The variables in the *expenditure module* are 'type of expenditure' (social benefits to individuals or households, social benefits to institutions, administration costs, transfers to other schemes, transfers to reserves, net expenditure for loaning assets or ground, other expenditures), 'purpose' (prevention, recovery, compensation), 'social security function ESSPROS', 'social security function ILO', 'social security function OECD', 'type of benefit' (in cash, reimbursements, cheques and vouchers, goods and services, re-routed social contributions), 'computing base' (fixed, according to the (last) income, according to another criterion, combination of income and other criterion), 'expenditure limits' (no limits, minimum, maximum, minimum- and maximum), 'contributions related' (yes or no), 'type of right' (personal right or derived right) and the amount of the expenditures paid and to be paid, taking into account possible arrears, advances, reclamations, refusals, reductions and deductions.

---

[4] CLA = collective labour agreement

The variables in the *financing module* are 'type of financing' (actual employers' social contributions, imputed employers' social contributions, social contributions by employees, social contributions by self-employed, social contributions by public servants, social contributions by beneficiaries, social contributions by other protected persons, re-routed social contributions, earmarked taxes, general revenue, transfers from other schemes, transfers from reserves, income from investments, property income, other receipts, transfer from RSZ Global Management, transfer from RSVZ Global Management)[5], 'charging base' (fixed, employer's total labour costs, employee's wage, self-employed person's income, public servant's wage, benefits, contributions, household income), 'financing limits' (no limits, minimum, maximum, minimum and maximum) en the amount of the receipts cashed or to be cashed, taking into account possible arrears, advances, reclamations, refusals, reductions and deductions.

In the *individual matrix* (Berghman & Nijs, 2005, pp. 63-111) individuals are the unit of observation. They are scored on several variables, divided into three modules. In the person module some personal features are recorded. The family module aims at characterising nuclear families and the household module does the same for private households.[6] The data in the individual matrix have to refer to the situation at the last day of the quarter.

The variables in the *person module* are 'sex', 'date of birth', 'country of birth', 'country of citizenship', 'usual place of residence' (postal code of the community),'current educational situation' (school going, further training or education, currently no education or training) 'level of education' (no education, primary education, lower secondary education labour oriented, lower secondary education oriented to upper secondary education, upper secondary education labour oriented, upper secondary education oriented to tertiary education, post-secondary non-tertiary education, tertiary education first stage practical orientation, tertiary education first stage theoretical orientation, tertiary education second stage), 'activity status' (employed, unemployed, student, housewife/man, (early) retirement beneficiary, disabled person, other, person aged under 15), 'number of jobs', 'employment status' (worker private

---

[5] RSZ = National Office for Social Security, RSVZ = National Institute for Social Security of the Self-Employed
[6] There are two methods for defining a private household. The *household dwelling concept* defines a private household as the persons occupying the same housing unit. The *housekeeping concept* defines a private household as the persons who combine to occupy a housing unit and to provide themselves with food and possibly other essentials for living. They may pool their incomes to a greater or lesser extent (Eurostat, 1999, p. 13; United Nations, 1998, p. 44). Nuclear families are two or more persons within a private household who are related as partners or as parent and child. Thus, a family comprises married, registered or consensual cohabiting couples with or without children and lone parents with children (United Nations, 1998, p. 45).

sector, employee private sector, worker public sector, employee public sector, public servant public sector, helper of own-account worker, own-account worker with employees, own-account worker without employees, member of producer's cooperatives, contributing family worker, workers not classifiable by status), 'occupation' (ISCO-88-COM classification), 'branch of activity' (NACE-BEL-classification), 'sector' (corporations, general government, households, non-profit institutions serving households, units outside Belgium), 'time usually worked' (in hours per week), 'income from employment' (amount in €),'civil status' (single, registered partnership, married, annulled marriage, putative marriage, actually separated, separated from table and bed, divorced, widowed), 'de facto living arrangement' (married, registered partnership, consensual partnership, not living with a partner), 'type of household' (nuclear family household, living alone, living with relatives, living with non-relatives, educational institution, health care institution, institution for elderly, military institution, religious institution, other institutional household, other household), 'household status' (single, spouse, registered partner, consensual partner, own child, foster/adoption/stepchild, spouse/registered/consensual partner of child, own sibling, partner's sibling, grand child, parent, foster/adoption/stepparent, partner's parent, grand parent, partner's grand parent, other relative, partner's other relative, spouse/registered/consensual partner of own or partner's relative, non-relative), 'family status' (spouse, registered partner, consensual partner, lone parent, child of both partners, child of either partner, child of lone parent), 'widow's status' (orphan of either parent, orphan of both parents, widowed after marriage, widowed after registered partnership, widowed after consensual partnership), 'identification number deceased parent/partner', 'date of death parent/partner', 'date of death', 'household referent' (person is household referent, person is not household referent, person does not live in a private household), 'family referent' (person is family referent, person is not family referent, person does not live in a nuclear family), 'identification number household referent' and 'identification number family referent'.

The variables in the *family module* are 'type of family' (married couple without children, registered cohabiting couple without children, factual cohabiting couple without children, married couple with at least one child, registered cohabiting couple with at least one child, factual cohabiting couple with at least one child, lone parent with at least one child), 'reconstituted family' (yes or no), 'family size' (total amount of family members),'amount of members younger then 18', 'amount of members between 18 and 25', 'amount of members of 25 and older', 'amount of economically active members', 'amount of members above pension age', 'identification number new family member', 'identification number deceased family member', 'identification number member that moved to another family' and 'identification number member that left the family for another reason'.

The variables in the *household module* are 'type of private household' (one-person household, multi-person non-family household, one-family household with children in the household, one-family household without children in the household, two- or more-family household with children in the household, two- or more-family household without children in the household), 'household size' (total amount of household members),'amount of members younger then 18', 'amount of members between 18 and 25', 'amount of members of 25 and older', 'amount of economically active members', 'amount of members above pension age', 'tenure status' (a member of the household is owner of the housing unit, a member of the household is tenant of the housing unit, household occupying the housing unit under some other form of tenure), 'type of ownership' (owner-occupied dwelling, dwelling in private ownership, dwelling owned by the local or central government, dwelling owned by housing non-profit associations, other types of ownership), 'disposable household income' (amount in €), 'saving household' (saving or non-saving household), 'identification number new household member', 'identification number deceased household member', 'identification number member that moved to another household' and 'identification number member that left the household for another reason'.

The *connecting matrix* (Berghman & Nijs, 2005, pp. 112-120) relates a person to its relevant social security products. The four central relations are 'protected person', 'beneficiary', 'contributor' and 'person at charge'. The relation 'protected person' relates individuals to expenditure measures that they would receive if a certain risk or need occurs. The relation 'beneficiary' relates individuals to expenditure measures they currently receive. The relation 'contributor' relates individuals to financing measures that imply a personal contribution. Finally, the relation 'person at charge' identifies which individual is person at charge of which other individual according to the various meanings of that concept. The connecting matrix has to capture all possible social security positions a person took in during the quarter. The relation 'beneficiary' implies that for each individual all social security products he/she benefited from during the quarter have to be listed (also when these were not yet allocated and also if they are not rewarded due to refusal or suspension). This information has to be registered for the person to whom the benefit is allocated. Along with that also the identification number of the person who opens the right to that provision because of his/her current or past activity or social status and the identification number of the person for whom the allocation takes place have to be recorded. Furthermore, the start and end date of the allocation period and the amount of the benefit have to be indicated, taking into account possible refusals or reductions. Finally, the personal scope of application has to be coded as well, i.e. the precise code for the condition that renders the person protected for or that makes him/her beneficiary of the social security product at hand. The relation 'contributor' relates individuals to financing measures, to wit to social security products that are part of social security receipts and that can be linked to an individual on account of whom that payment is due (also when these were not yet paid and also if they are not to be paid due to reductions and exemptions). Again, the code of the personal scope of application of the contributions, the start and end date of the contribution period and the amounts have to be registered, taking into account possible exemptions, reductions and deductions.

3. Practical elaboration

The main aim of the research project is to develop a harmonised and consolidated picture of the Belgian population according to social security status and to build a database that contains information that is relevant for theoretical and policy research and that is in compliance with the international data requirements regarding social protection statistics. The social security matrix is an instrument that points out which information is needed to provide the demanded picture on Belgian social security and the social security population. After conceptualising the social security matrix and its methodological issues, it is of utmost importance that those concepts and ideas are validated. This validity test implies confronting the social security concept matrix with reality. The theoretical framework of the matrix has to be filled in with the required data. The matrix must function as the guiding principle for optimising the Belgian statistical arsenal with respect to social security in the long term. In the short term, however, we can already try to fit in the existing data into the outline of the matrix.

As for the product matrix, this means that all social security products (i.e. each social protection measure as well as each part of it disaggregated down to the individual level) have to be inventoried and scored on the variables of the product matrix. By mutual agreement with the FPS Social Security, the researchers chose in the first instance to focus on the schemes for inability to work. This resulted in several Excel-files named the 'Product Matrix for Belgian Social Security' (PROMAT).

For implementing the individual matrix and the connecting matrix, the researchers investigated which data are needed to do so and where these data are available. Thereupon and in collaboration with the research project AG/01/086 'Social security microsimulation model' a demand for coded individual social data was drawn up and submitted by the FPS Social Security at the Crossroads Bank for Social Security. By means of the data rendered to our disposal, a database with various information on the individual level was constructed: the 'Person Matrix for Belgian Social Security' (PERMAT).

3.1. PROMAT: inventory of social security products

Already in the Harmonisation Project the product matrix was put to a first validity test (by means of the social protection measures regarding inability to work to be more exact). During the Validation Project it was further validated with respect to some crucial aspects. First of all, the variables were brought in compliance with new international data demands. Next to that, the matrix was validated for the financial measures. This resulted in some adjustments with respect to some variables in the product matrix. Other modifications were made by the occasion of validating the matrix for second and third pillar provisions.[7] Furthermore, the longitudinal aspect of the matrix was designed in more detail and the product matrix for inability to work was updated. It covers now the period from the first quarter of 1999 to the last quarter of 2004. Finally, also the personal scope of application was inventoried.

The inventory and coding of social security products is materialised in Excel files. For the moment being, the relevant amounts are not yet included since they are not available as required. The way in which the institutions of social security administrate and manage their data is not tuned to the outline of the matrix.

---

[7] The validity of the matrix with respect to second and third pillar provisions is tested within the course of three other projects by our research team: SO/01/006 'Invisible pensions in Belgium', AG/01/084 'Supplementary social provisions 2nd and 3rd pillars: preparation of a database' and S0/01/61 'The implications of labour market flexibility for social security'.

3.2. PERMAT: database at the personal level

For the variables in the individual matrix and the relations in the connecting matrix, the researchers investigated which data from which data files are needed in order to actually fill in some parts of the social security matrix. Soon it became clear that not all required data are available and that some other data are not available in the right form. For the data that did could be used, the FPS Social Security submitted a data demand at the Crossroads Bank for Social Security. It concerned a demand for coded individual social data, which will be used not only in the project AG/01/070 'Validation of the individual part of the social security concept matrix' but also in the project AG/01/086 'Social security microsimulation model'.

From the National Register a sample of 100.000 was taken randomly. Next, for those persons also the other members of the private household were gathered in the sample as well. In this manner, the final sample encompasses 305.019 persons. In order to collect the demanded data, several data sources from diverse institutions for social security[8] have to be called upon and it was agreed that this would be done in different phases. In June 2004 we received the first wave of data, namely from the National Register and from the existing Datawarehouse Labour Market at the Crossroads Bank: LATG data of the RSZ, STATPLUS data of the RSZPPO, ARZA data of the RSVZ, benefits data of the RVA, data of the Medical Council for Invalidity of the RIZIV and data on child benefits from the RKW and the RSVZ. The benefits data of the RIZIV and the data from the Pension Register were delivered in January 2005, followed by the data of CIMIRe in May 2005. The remaining demanded data are not yet delivered to or processed in the Datawarehouse and hence could not be handed over to us at the moment being. It concerns data of the FAO from its central database, data of the FBZ, data on sickness benefits of the IMA, fiscal data of the NIS, data of the FAO from its financial database, data of the RVP and the AP, data on benefits to handicapped persons of the FPS Social Security and census data from the General Socio-Economic Survey 2001 of the NIS.

---

[8] RSZ = National Office for Social Security, RSZPPO = National Office for Social Security of the provincial and local government administrations, RSVZ = National Institute for the Social Security of the Self-Employed, RVA = National Office for Employment Measures, RIZIV = National Institute for Sickness and Invalidity Insurance, RKW = National Office for Employees' Child Benefits, CIMIRe = Multisectoral Individual Account, FAO = Fund for Work-Related Accidents, FBZ = Fund for Occupational Diseases, IMA = agency covering the various national health insurance services, NIS = Official Belgian Statistical Institution, RVP = National Office for Pensions, AP = Pension Administration of the FPS Social Security.

The data we already have at our disposition are used mainly to fill in the individual matrix. The precise connection between individuals and social security products (as outlined in the connecting matrix) is not yet operable. The data available are not detailed enough to do so and the way in which the social security institutions administrate and manage their data are not tuned to the outline of the matrix. Still, by means of the data we already have at our disposal, a database on the personal level was constructed. Due to lack of time, the data of the Pension Register and CIMIRe could not be processed and hence these are not part of the database.

Of all 305.019 individuals in the sample, information from the National Register (situation on January 1st 2002) was delivered. Their unique social security identification number was converted by the Crossroads Bank into a coded number and these coded numbers are also entered in the other data sources (for it concerns a coupling of administrative data by means of the unique social security identification number of individuals). Next to the National Register, only the data of the Medical Council for Invalidity of the RIZIV and the derived variables from the Datawarehouse Labour Market have individuals as the unit of observation (one observation line per person). In all other data sources individuals may appear one more than one observation line. This is the case for the RSZ data, the RSZPPO data, the RSVZ data, the RVA data, the benefits data of the RIZIV and the child benefit data of the RKW and the RSVZ. These data were aggregated in such a way that the data in the new database are all on the personal level (all variables appear on one observation line per person). For practical reason, the database is split up into several SAS data files by theme and by quarter: one file with variables of the *person, household and family module* (only for the fourth quarter of 2001), eight files with data on people's *labour market situation* (some variables only for the four quarters of 2000 and some other variables also for the four quarters of 2001), eight files with data one the *personal scope of application, contributions and contribution reductions* (idem), eight files with data on *inability to work* (idem), eight files with data on *unemployment* (idem) and four files with data on *child benefits* (only for the four quarters of 2001). Obviously different variables for different quarters from different files can be assembled in order to run the analyses one wants to perform.

4. Conclusion and future aspects

At the suggestion of the Crossroads Bank for Social Security, it was agreed that the demanded data would not be retrieved and delivered just for single use, but that they firstly would be valuably integrated in the existing Datawarehouse of the Crossroads Bank (which covered until then mainly information concerning the labour market). Thus, the former Datawarehouse Labour Market is to be supplemented gradually with new data files and in this way, it moves to a Datawarehouse Labour Market and Social Protection.

Using administrative data for statistical purposes requires that these data are harmonised, that they are comparable and adjusted as regards content and that they are utilised and interpret correctly. The social security matrix has to be viewed as an important guide in harmonising and improving administrative data on social security, so that these can be used for innovating statistics and policy research. It is a general concept matrix, a comprehensive framework that shows which information in which way is needed to picture Belgian social security in all its aspects in a consistent and structured manner. The product matrix provides an extensive and at the same time detailed description of the social security system. Moreover, by relating this to the individual matrix, also the Belgian population in terms of social security categories becomes visible and one can move from statistics on expenditures and receipts to statistics on contributors and beneficiaries. Important socio-economic trends and their impact on social protection can be researched.

Bibliography

Adema, W., & Einerhand, M. (1998). *The growing role of private social benefits. Labour Market and Social Policy. Occasional Paper No. 32*. Paris: OECD.

Adema, W. (2001). *Net Social Expenditure. Second Edition. Labour Market and Social Policy. Occasional Paper No. 52*. Paris: OECD.

Berghman, J. (1986). *De onzichtbare sociale zekerheid*. Deventer: Kluwer.

Berghman, J., & Nijs, K. (2003). Social governance: een niet vervuld perspectief. In Instituut voor Sociaal Recht (Red.), *Sociale bescherming op nieuwe paden. Liber Memorialis Béatrice Van Buggenhout* (pp. 449-462). Leuven: Universitaire Pers.

Berghman, J., Nijs, K., & Nachtergaele, W. (2004). The development of a new statistical instrument to survey social protection in Belgium. In European Communities (Red.), *23rd CEIES seminar. Social Protection Statistics* (pp. 177-186). Luxembourg: Office for Official Publications of the European Communities.

Berghman, J., Booghmans, A.-M., Boos, C., Nachtergaele, W., Verhalle, I., & Vleminckx, K. (2004). *Harmonisatie van de statistische concepten in de sociale zekerheid. Agora-project AG/01/039. Eindrapport december 2001*. Leuven: Departement Sociologie.

Berghman, J, & Nijs, K. (2004). *Het ontwikkelen en valideren van een globale conceptmatrix van de sociale zekerheid. Agora-project AG/01/070. Tussentijds rapport april 2003*. Leuven: Departement Sociologie.

Berghman, J, & Nijs, K. (2004). *Het ontwikkelen en valideren van een globale conceptmatrix van de sociale zekerheid. Bijlagen. Agora-project AG/01/070. Tussentijds rapport april 2003*. Leuven: Departement Sociologie.

Berghman, J, & Nijs, K. (2005). *Handleiding voor de implementatie van de sociale zekerheidsmatrix*. Leuven: Centrum voor Sociologisch Onderzoek.

Eurostat (1995). *European System of Accounts ESA 1995*. Luxembourg: Office for Official Publications of the European Communities.

Eurostat (1996). *Esspros Manual*. Luxembourg: Office for Official Publications of the European Communities.

Eurostat (1999). *Guidelines and Table Programme for the Community Programme of Population and Housing Censuses in 2001. Volume 1: Guidelines*. Luxembourg: Office for Official Publications of the European Communities.

Eurostat (2000). *Harmonisation of recommended core units, variables and classifications*. Luxembourg: Eurostat Directorate E.

Eurostat (2001). *Draft questionnaire for the EU-SILC pilot experiment*. Luxembourg: Eurostat Unit E2.

Expert Group on Household Income Statistics (2001). *Final Report and Recommendations*. Ottawa: United Nations.

ILO (29.10.1982). *Resolution concerning statistics of the economically active population, employment, unemployment and underemployment, adopted by the Thirteenth International Conference of Labour Statisticians* [WWW]. ILO: http://www.ilo.org/public/english/bureau/stat/res/ecacpop.htm [10.10.2002].

ILO (28.01.1993). *Resolution concerning the International Classification of Status in Employment, adopted by the Fifteenth International Conference of Labour Statisticians* [WWW]. ILO: http://www.ilo.org/public/english/bureau/stat/res/isce.htm [03.05.2002].

ILO (1994). *Cost of Social Security 1994-1996: Nineteenth International Inquiry* [WWW]. ILO: http://www.ilo.org/public/english/protection/socsec/publ/css [?].

ILO (2004). *ILO Social Security Inquiry Manual. First Inquiry 2004/2005. Draft version 29 June 2004*. Geneva: International Labour Office. Social Protection – Financial, Actuarial and Statistical Services Branch.

OECD (2001). *1980-1998. 20 years of social expenditure. The OECD Database*. Paris: OECD.

UNESCO (1997). *International Standard Classification of Education ISCED 1997*.

United Nations (1998). *Statistical Standards and Studies No.49. Recommendations for the 2000 Censuses of Population and Housing in the ECE Region jointly prepared by the United Nations Economic Commission for Europe and the Statistical Office of the European Communities*. New York: United Nations.