

Max-Planck-Institut für demografische Forschung
Max Planck Institute for Demographic Research
Doberaner Strasse 114 · D-18057 Rostock · GERMANY
Tel +49 (0) 3 81 20 81 - 0; Fax +49 (0) 3 81 20 81 - 202;
<http://www.demogr.mpg.de>

MPIDR WORKING PAPER WP 2002-010
FEBRUARY 2002

**Integrated Information System
for Demographic Statistics
'ESGRAON-TDS' in Bulgaria**

Iliana Kohler (iliana.kohler@demogr.mpg.de)
Jordan Kaltchev (ikaltchev@nsi.bg)
Mariana Dimova (mdimova@nsi.bg)

This working paper has been approved for release by: James W. Vaupel (jwv@demogr.mpg.de)
Head of the Laboratory of Survival and Longevity.

© Copyright is held by the authors.

Working papers of the Max Planck Institute for Demographic Research receive only limited review.
Views or opinions expressed in working papers are attributable to the authors and do not necessarily
reflect those of the Institute.

Integrated Information System for Demographic Statistics ‘ESGRAON–TDS’ in Bulgaria

Iliana Kohler* Jordan Kaltchev† Mariana Dimova‡§

February 25, 2002

Abstract

In this paper we describe the Bulgarian statistical system that provides unique research possibilities for demographic analyses. In particular, we review the methodology, structure and informational flows for population registration and the registration of vital events in Bulgaria. In addition, we describe the main elements of the most recent Bulgarian census from March 1, 2001. In the last part of the paper, we pay special attention to data protection and data confidentiality issues in Bulgaria.

1 Introduction

Population data on individual level are vastly desirable as they usually provide the most extensive social, economic, health and other vital information for studying demographic events. Nevertheless, few countries are able to supply large-scale, individual-level statistical information for the entire population. The Nordic countries lead the world in this area: demographic statistics can be obtained from population-based registers with national coverage (see for example Frank 2000; Knudsen and Murphy 1999; Petersen 2000; Andersen et al. 1999; Eurostat/Statistics Denmark 1995 on registers as a data source for demographic and epidemiological analyses in Denmark, or Statistiska centralbyrån 2001 on Swedish registers).

In the Central and Eastern European countries, however, such large-scale, individual-level data are still scarce. Bulgaria represents an exception. In 1978, Bulgaria introduced

*Research Scientist, Max Planck Institute for Demographic Research, Doberaner Str. 114, 18057 Rostock, Germany. *Tel:* +49-381-2081-125, *Fax:* +49-381-2081-425, *Email:* iliana.kohler@demogr.mpg.de.

†Head of Department ‘Population, residential areas and GIS’, National Statistical Institute of Bulgaria, P. Volov Str. 2, 1504 Sofia, Bulgaria. *Tel:* +359-2-9857-2616, *Email:* IKaltchev@nsi.bg.

‡Head of Department ‘Data processing of Demographic Information’, National Statistical Institute of Bulgaria, P. Volov Str. 2, 1504 Sofia, Bulgaria. *Email:* MDimova@nsi.bg.

§The authors have equally contributed to this paper. For correspondence, please contact Iliana Kohler.

a unified system for population registration and administrative service to the population called ESGRAON. This administrative system mandates that all individuals residing in the country are assigned a unique personal identification number that is used as an identifier in almost all administrative affairs. This administrative structure allows one to link individual-level information from various data sources and provides unique research possibilities.

This potential of the ESGRAON system has been explored for the first time in dissertation research by Iliana Kohler (Kohler 2001). In particular, the author uses a mortality dataset covering the entire Bulgarian population that was obtained from a link between the Bulgarian census from December 4, 1992 and the death certificates for the following year. The link between these two types of demographic information for the Bulgarian population—the death certificates as a part of the vital statistics and the population census—was possible because of the personal identification number which provided a unique identifier for the link. The linkage procedure was carried out at the National Statistical Institute of Bulgaria under consideration of all data protection requirements. The link is of very high quality: in total 93.7 per cent of all death certificates that were issued during the period between December 5, 1992 and December 31, 1993 in Bulgaria have been linked to the census records of the respective individuals. The constructed mortality dataset represents the first dataset of this type in Bulgaria. Moreover, it is also unique for other Central and Eastern European countries. Similar possibilities for data links also exist for fertility.

This paper provides important background information by describing the structure of Bulgarian demographic statistics and drawing attention to various characteristics that provide unique research possibilities in Eastern Europe.

Four main sources of information of the Bulgarian population and the demographic events in the country can be distinguished:

1. The Unified System for Population Registration and Administrative Service to the Population—ESGRAON (in Bulgarian: ‘Edinna Systema za Grashdanska Registracija i Administrativno Obslushvane na Naselenieto’);
2. The vital statistics system—TDS (in Bulgarian: ‘Tekusta Demographska Statistika’);
3. Population censuses;
4. Specialized studies focusing and investigating specific demographic processes/events in the Bulgarian population such as family and fertility surveys, health surveys, migration and labor market studies.

In the following sections, we focus primarily on the first three sources of demographic information in Bulgaria and describe their structure and the data flow process. We skip the presentation of the various specialized demographic studies in Bulgaria as the description

of their focus and methodology is beyond the purpose of this paper. In addition, in the last section of this paper we discuss the data protection and data confidentiality issues in Bulgaria.

2 Integrated Information System for Demographic Statistics ‘ESGRAON–TDS’

The integrated information system for demographic statistics, ‘ESGRAON–TDS,’ consists of two basic components: *a)* the unified system for population registration and administrative service to the population (ESGRAON), and, *b)* vital statistics (TDS). These components are administered separately. The corresponding departments at the Ministry of Regional Development and Public Works are responsible for the management and administration of the citizens’ civil registration (i.e., the ESGRAON component), while the National Statistical Institute (NSI) processes and publishes the vital statistics (i.e., the TDS component). However, even though the two components and their administration are independent of each other, the respective departments of the two institutions—the Ministry of Regional Development and Public Works and the National Statistical Institute—collaborate and coordinate with each other in the development of the documents for the registration of the vital events, the development of informational standards (i.e., classifiers and nomenclatures) and the exchange of information between the two institutions.

2.1 Informational Basis of the System

The integration of the two major components, ESGRAON and TDS, into a large information system for demographic statistics in Bulgaria requires a uniform registration and administration of the individual personal information/records for the country’s citizens and the demographic events registered to them. Thus, the ‘ESGRAON-TDS’ system is based on:

1. standardized documents for the identification and registration of the civic status of the population,
2. uniform national classifiers and nomenclatures, and
3. informational blocks about individuals and events on the municipal and national level that are stored as original documents, computer files and data bases.

The standardized documents for the identification and registration of the individual’s civic status include birth, death, and marriage certificates. These are official documents and are stored in the respective municipal Civil Registration Offices for a period of 130 years.

The civic and address registration of the individuals is based on notifications for birth, marriage, divorce and changes in the address registration. The notification documents

include the relevant information, on the basis of which the corresponding official certificates are issued. However, in contrast to the official certificates, the notification documents do not have a juridical power; they simply state that a certain event happened. In addition, these notifications include information required for demographic statistics, information about the officials issuing the document, stamps, etc. The notification documents are stored for a period of two years and are afterwards destroyed according to regulations for ensuring confidentiality of individual data. In the future, it is planned that the notification documents for birth and death will be stored according to the regulations and terms foreseen for the official birth and death certificates (i.e., 130 years).

The uniform national classifiers and nomenclatures are as follows:

1. Uniform classifiers of the administrative and territorial units in Bulgaria [in Bulgarian: ‘Edinen klasifikator na administrativno-teritorialni i teritorialni edinici v stranata’, abbreviated as EKATTE]. These classifiers are managed and regularly updated by the National Statistical Institute of Bulgaria.
2. National classifier of traffic arteries [in Bulgarian: ‘Nacionalen klasifikator na patnite arterij v stranata’ abbreviated as KPA], national classifier of addresses and sites [in Bulgarian: ‘Nacionalen klasifikator na adresite i dvornite mesta v stranata’ abbreviated as KAD]. Both classifiers are managed and updated by the Ministry of Regional Development and Public Works.
3. Other classifiers, e.g. of countries, or the International Classification of Diseases (at present, the ICD-9 revision is used in Bulgaria).
4. Categorization of indicators that are part of the integrated information system for demographic statistics such as education, marital status, reasons for changes in the address registration, civil status, and others.

2.2 Informational Coverage of the System

The integrated information system for demographic statistics ‘ESGRAON-TDS’ provides the following information:

1. The ESGRAON component supplies information about the civic status of those individuals who are subject to registration in the ESGRAON system. These are all Bulgarian citizens and foreigners with a different status who are residing in the country for an extended period (according to the Geneva Convention and the law for foreigners in Bulgaria). The ESGRAON system provides full identification of the individuals as it includes information on the full names of the individuals subject to registration in the system, their unique personal identification number, and their civil status and address registration (permanent and present address). In addition, there is information about the number, place, date and office issuing the respective

official document certifying the civil status of the person as well as other bureaucratic information.

2. The TDS component (the vital statistics) includes information about demographic events including birth, death, marriage, divorce, and migration within the country. For the correct management and data flow of the demographic statistics it is required that each demographic event be coded with a serial number and information about when and where this demographic event has occurred. In addition, information about the socio-demographic characteristics of all individuals who have ‘participated’ in the corresponding demographic event is supplied: that is, their personal identification number, sex, age, marital status, degree of education, place of residence, etc.

2.3 Organization of the ‘ESGRAON-TDS’ System

The organization and function of the ‘ESGRAON-TDS’ information system for demographic statistics is closely related to the regional and administrative structure of the country. Bulgaria is currently divided into 263 municipalities that are governed by a major and a municipality council. Both, the major and the council, are directly elected by the residents of the municipality. Within these municipalities, one or more residential areas (depending on their size) can build a city/town council that is headed by a directly elected major.

Adjoining municipalities are grouped into administrative districts called ‘oblast’. These 28 administrative districts in Bulgaria are governed by the district administration. The district administrations are local decentralized authorities that are, however, part of the national executive power.

The ‘ESGRAON-TDS’ system is based on two types of documents that are officially issued:

1. *Notification documents* issued by the corresponding official institutions such as health care institutions, local administrative authorities, and courts.
2. *Certificates* issued by the local administrative authorities, who have juridical power and certify the civic status of the person.

On the basis of the notification documents, the local administrations regularly update the local population registers. These registers contain information for all permanent residents in the residential area. The notification documents are afterwards sent to the ‘Civil Registration Office’ on the municipal level.

The municipal ‘Civil Registration Office’ controls the information included in the notification documents. In addition, it is also responsible for the coding and computerizing of this information as well as up-dating the municipal population registers. Afterwards, the computerized information, as well as the original notification documents, are sent to the district office for civil registration and administrative service to the population called

‘GRAO’ [in Bulgarian: ‘Grashdanska registracija i administrativno obslushvane’]. The latter is responsible for the correct data flow and exchange of information between the residential areas within the district or between different districts, when the event occurs in a different place than the permanent address of the individuals. The district office is also responsible for the quality control of the information. In case there are some irregularities in the individual data, the local authorities are notified and if necessary, the documents are sent back to the initial level. The district office also compiles and up-dates the population data for the entire district.

Part of this regularly up-dated information flows into the ESGRAON system on the district level. The same information is also entered into the ‘National Database Population,’ which is managed and up-dated by the main department for civil registration and administrative service to the population ‘GRAO’ at the Ministry for Regional Development and Public Works.

The computerized population information on the district level is also used to compile the vital statistics (TDS). The National Statistical Institute of Bulgaria is represented in each of the 28 districts by its own office located in the administrative center (town/city) of the district. The computerized files with the population information on the district level is sent regularly to the employees of the National Statistical Institute on a district level. These files include data on all demographic events that have occurred in the district within the last month, and are made available to the statistical bureaus until the 15th of the following month.

The methodology (instructions and regulations), the organization of the system and computerizing process of the information about the population and the demographic events are developed and administered by the different working groups from the main department ‘GRAO’ at the Ministry for Regional Development and Public Works with the assistance and collaboration of the National Statistical Institute. The processing of the data and its use is further carried out according to the requirements of the two informational systems.

The structure, functioning and data flows of the integrated information system for demographic statistics ‘ESGRAON–TDS’ is shown graphically in Figure 1.

2.4 The Personal Identification Number—EGN

The key element for the organization of the system for administrative service to the population (ESGRAON) is the unique personal identification number [in Bulgarian: ‘edinen grashdanski nomer’ abbreviated as EGN] that is assigned to each Bulgarian citizen by the district office ‘GRAO’.¹ The personal identification number is unique for each Bulgarian citizen and is used as a unique identifier in all administrative information systems in the

¹Foreign citizens who are permanent residents or have an official status as refugees in Bulgaria are also assigned a personal identification number. Foreigners with a different status from the one mentioned above, but who live in Bulgaria for a certain period of time, receive a special personal number for foreigners, which has a different structure.

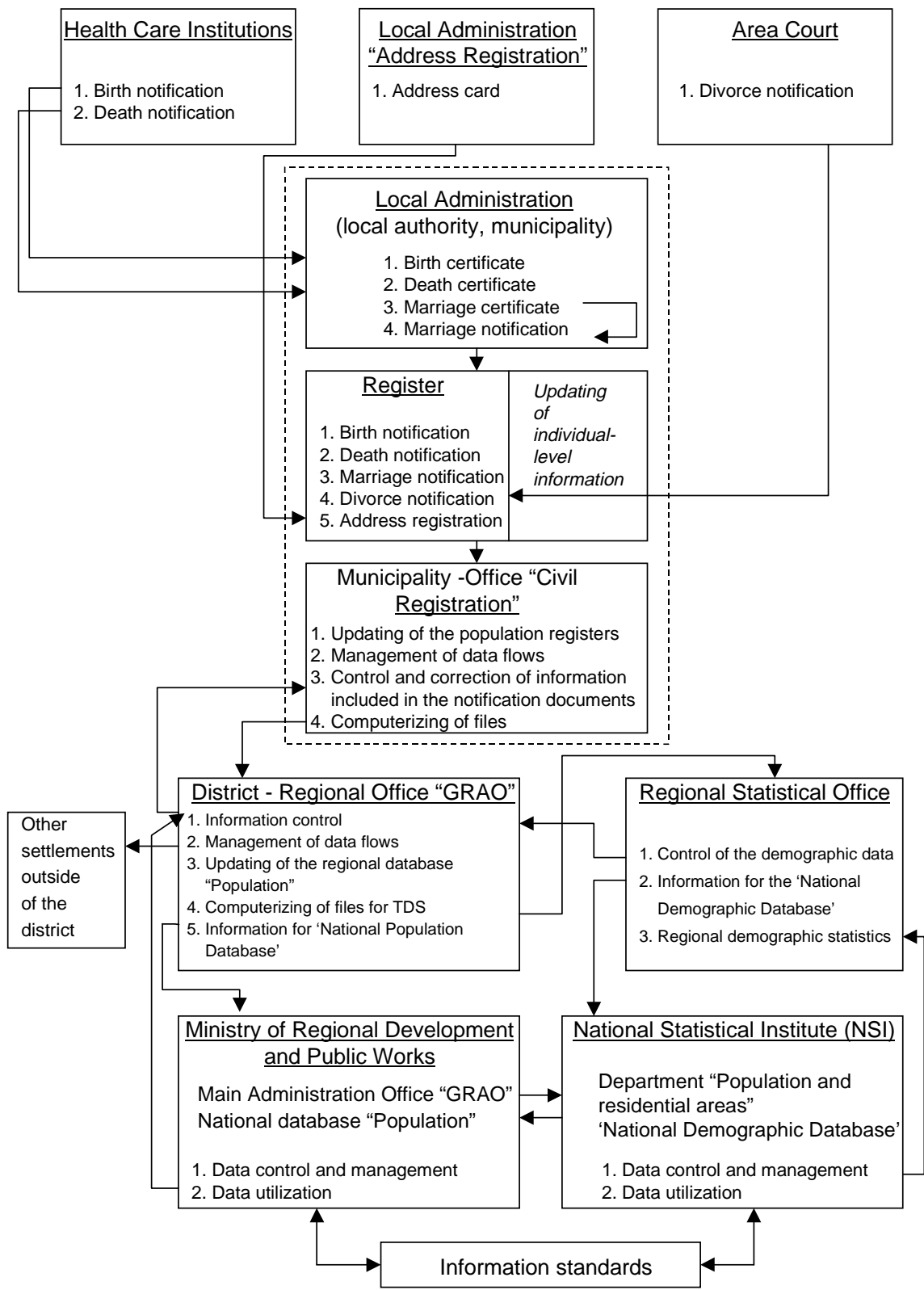


Figure 1: Structure and data flows of the integrated information system for demographic statistics 'ESGRAON-TDS'.

Structure of the personal identification number (EGN)

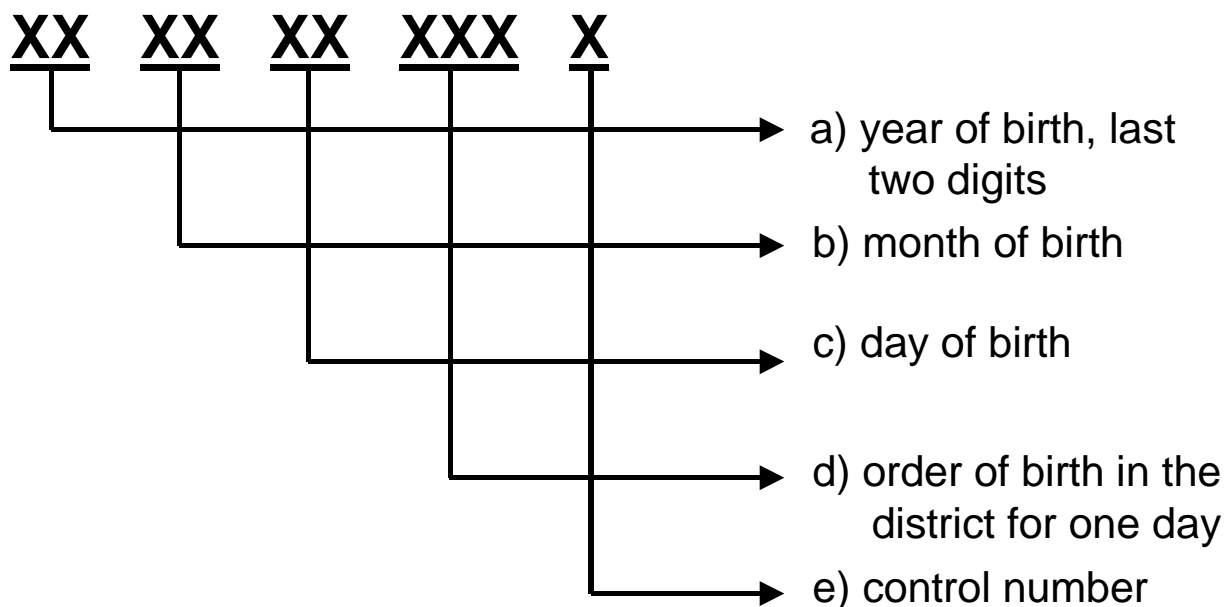


Figure 2: Structure of the 10-digit personal identification number (EGN) in Bulgaria.

country, in the national retirement and health systems, and in banks, etc. Each administrative information system that uses the EGN for a unique identification of individuals controls the validity of the personal identification numbers entered in their system by applying ‘Module 11’.

The personal identification number is a 10-digit number that is assigned to each newborn in the district where the birth occurred, or the district where the parents (or mother) have their permanent residence. The structure of the personal identification number is shown in Figure 2. The personal identification number (EGN) is based on the birth data of each individual. The first two digits give the year of birth, the next two digits correspond to the month of birth. The fifth and sixth digits match the day of birth, the following three numbers show the order of the birth in the district for that specific day, and the last number is a control number used to check whether the coded personal identification number is correct.

In order to avoid problems with duplication of numbers when people are born in different centuries, the structure of the personal identification number can be slightly modified. For all individuals born before 1900, the month of birth is coded as $month + 20$. For example, this means that if a person was born in the month of January in the 19th

century, his or her month of birth is coded as 21. If the person was born in December, the birth month code would be 32.

All people born in 2000 and afterwards are assigned a personal identification number in which the birth month is coded as $month + 40$. For example, if a person was born in January 2000, then the two digits for month of birth in the EGN number are coded 41. If the person was born in December 2000, then the month of birth in his/her EGN number is coded 52.

All individuals born between 1900 and 1999 have received a personal identification number in which the third and fourth digit correspond exactly to the month order without any modifications.

The sequence of the birth (the seventh to the ninth digit in the EGN number) corresponds to the order of birth for a given day in the district where the birth occurs. Each administrative district gets an allocation of numbers for a given day. Births are assigned a number in the set according to the sequence of birth on that day. Males are assigned even numbers and females are assigned odd numbers.

The tenth digit is a control number that results from an algorithm (the so-called ‘Module 11’) and is calculated on the basis of all other numbers in the EGN number.

Example 1 *Assume that a boy was born on the September 15, 2000, in the town Pleven. Assume that this is the fourth male birth on this day in the town. This boy will be assigned the following EGN number: 00 49 15 396 X, where 00 indicates that the boy was born in the year 2000. The next two digits 49 show that the boy was born in September ($month + 40$), and 15 shows the day of birth. The number 396 corresponds to the sequence of birth in the district. This district gets a quota of numbers from 390 to 429. The number 396 shows that this boy was the fourth birth in the district on this day (15th of September), and the even number shows that this was the fourth male birth in the district. The last digit, which we call ‘X’ is the control number.*

3 Vital Statistics (TDS)

The data on Bulgarian population and its socio-demographic characteristics are calculated on the basis of the Bulgarian censuses and the yearly vital statistics (TDS). The yearly population statistics are based on the natural change and in- and out-migration of the Bulgarian population.

The National Statistical Institute of Bulgaria is in charge of the regular implementation of censuses in the country and the compilation of the vital statistics. In particular, the ‘Department for Demography’ at the National Statistical Institute is responsible for the development of methodology, for compiling data and for the data processing of demographic statistics.

The main sources for the calculation of the natural change and migration in Bulgaria are the standardized documents included in the integrated information system ‘ESGRAON-TDS’. Each demographic event such as birth, death, marriage, divorce, or address change

is registered daily via notification documents in the ‘ESGRAON-TDS’ system. This data is first computerized by the municipal ‘Civil Registration Office’. Afterwards, the computerized demographic information is forwarded to the district offices of the National Statistical Institute that perform further checks for inconsistencies in the data. In case data inconsistency is detected, the information is corrected with the help and assistance of the employees at the district office ‘GRAO’. The controlled and corrected data is then sent to the main office of the National Statistical Institute in Sofia, where the data is entered in the ‘National Demographic Database’ (‘Nationalna Baza Demographija’). This procedure is carried out monthly.

3.1 Socio-demographic Indicators Included in the Vital Statistics

The notification documents contain information about the event itself as well as each person involved in the event. In particular, the following sociodemographic indicators are required for the compilation of the vital statistics and the maintenance of the ‘National Demographic Database’ at the National Statistical Institute:

1. Birth notification document

- (a) Personal identification number of the new-born
- (b) Sex
- (c) Date of birth
- (d) Citizenship
- (e) Whether the child is born in conjugal union or not
- (f) Born alive or still-born
- (g) If still-born, reasons for the still-birth
- (h) Height and weight
- (i) Duration of pregnancy
- (j) Permanent and present address of the new-born²
- (k) *Father’s data* (optional for non-married): personal identification number, education, citizenship
- (l) *Mother’s data*: personal identification number, education, citizenship, order of marriage, date of marriage, parity, number of surviving children, date of last birth

2. Death notification document

- (a) Personal identification number of the dead person

²Permanent address is the address, where the person is registered in the population registers. Present address is the address where the person lives at the moment, when the event occurs.

- (b) Sex
- (c) Date of birth (for children younger than one month, the exact hour of death is also required)
- (d) Citizenship
- (e) Cause of death (according to the ICD 9)
- (f) Permanent and present address
- (g) Marital status
- (h) Education

3. Marriage notification document

- (a) Date of marriage
- (b) Personal data required for both husband and wife:
 - i. Personal identification number
 - ii. Date of birth
 - iii. Citizenship
 - iv. Permanent and present address
 - v. Education
 - vi. Marital status before current marriage
 - vii. Number of children under age of 18 years
 - viii. Order of marriage

4. Divorce notification document

- (a) Date of marriage
- (b) Date of divorce
- (c) Number of surviving children born within this marriage
- (d) Reason for divorce
- (e) Recrimination for divorce
- (f) Personal data required for both husband and wife:
 - i. Personal identification number
 - ii. Date of birth
 - iii. Citizenship
 - iv. Permanent and present address
 - v. Education
 - vi. Marital status before the current marriage
 - vii. Order of marriage

viii. Order of divorce

5. **Address registration card**³

6. Personal identification number
7. Statute of the person
8. Present address
9. Previous address
10. Date of address registration
11. Education
12. Marital status
13. Reason for the stay
14. Duration of the stay
15. Children under age 18 who live with the registered person (their personal identification number is also required)

3.2 Methodological Criteria for the Organization of Vital Statistics

The main criteria for registration of demographic events is to refer their occurrence to the place of permanent residence of the persons involved in the event. Accordingly, births are registered in the place of permanent residence of the parents or the mother if she is single; deaths are recorded in the place where the dead person had his/her permanent address; marriages are usually registered in the place, where the husband has his permanent address; divorces are recorded in the place where both spouses lived permanently before separating.

The annual population data is based on calculations from the censuses and the intercensal estimates obtained from the vital statistics using data on the natural movement of the population and migration. Until 1964, the population data was calculated on the basis of the present population, but since 1965 the population refers to the permanent population (additional information about the calculation of the population can be found in National Statistical Institute of Bulgaria 1999, 2000).

3.3 National Demographic Database

The National Demographic Database is based on the 'Entity Relationship Model.' The database consists of three different levels of information that are connected to each other.

³This data is usually used for studies on migration within the country.

1. **Level One – Microdata.** This level contains individual records for each demographic event that occurs during a calendar year in Bulgaria. The microdata includes information about individuals and demographic events that are subject to the demographic statistics. The main sources of information for this data are the population censuses, the vital demographic statistics and other specialized studies or representative surveys. The data is computerized and stored as tables, in which each row represents a single individual case and each column represents certain socio-demographic characteristics that are subject to observation in the demographic statistics.

Currently, the microdata includes yearly files with information about the demographic events (births, deaths, marriages, divorces, and changes in the address registration) during the period 1995–1999 for the entire Bulgarian population. These files consist of individual records for each individual who has experienced one or more demographic event(s). In addition, the microdata includes all individual records from the 1992 population census.

The micro-level database is a relational database and uses the personal identification number (EGN) as a unique identifier (a key variable) for each individual record.

Access to the individual records is extremely limited in order to meet the data confidentiality requirements. The data is accessible to a restricted number of people only for updating the information, correction or control of the records, or for performing some special data linkages.

2. **Level Two – Macrodata.** The macrodata are aggregated files based on the individual-level records. The files include aggregated information about the number and structure of the Bulgarian population on the level of residential areas that represent the lowest level of territorial division. The macrodata are compiled on the basis of the microdata, yearly vital statistics and the population censuses.
3. **Level Three – Metadata.** This level includes a detailed description of the micro- and macro-level data in a semantic, structural and statistical aspect. It includes a detailed description of the data processing, storage and usage of the data. This information is complemented by a description of the general structure of the database, the methodology of specialized representative studies and surveys carried out in Bulgaria, possible information outputs from the database, legislative regulations and other important rules that regulate the use, construction, development and functioning of the database.

3.4 Publication of Demographic Data

In the first half of each calendar year, the National Statistical Institute publishes the specialized ‘Population’ bulletin that contains recent data about the Bulgarian population and its main socio-demographic characteristics at the national and regional (district) levels.

The bulletin also contains information on the natural change and migration within the country.

In addition, the National Statistical Institute publishes the yearly publication ‘Demographic Processes’ that describes the natural change of the Bulgarian population and the central demographic processes such as births, deaths, marriage, divorce and internal migration. Moreover, aggregated demographic data is also published in the ‘Statistical Yearbooks’ of Bulgaria. In the last years, the National Statistical Institute started to present aggregate-level data on its webpage (<http://www.nsi.bg>, the information can be found in Bulgarian and English language). Information about various publications of the National Statistical Institute related to population and other issues can be also found on the homepage.

4 Population Censuses

The foundation of the Bulgarian statistical system was set up shortly after the liberation from the Turkish empire in 1878. The first statistical department in the new independent state was founded per a Princely Decree No. 296 on June 25, 1880. This statistical department was set up in the Ministry of Finance. In 1881, which is officially considered the year of birth of Bulgarian statistics, a new independent statistical bureau was founded by Princely Decree No. 712 from August 19, 1881 (National Statistical Institute 2001).⁴

The establishment of the new statistical institution coincided also with the first population census conducted in Bulgaria after the country became independent from the Turkish Empire. The first population census in the Principality of Bulgaria was conducted on December 31, 1880. This census was followed by two censuses in Eastern Roumelia (present-day South of Bulgaria), which was an independent Bulgarian province until 1885 when it was united with the Principality. The task of these censuses was to collect information about the number and main socio-demographic characteristics of the population such as sex, age, marital status, residence, religion, nationality, mother tongue, literacy, etc.

A special achievement of the statistical practice at that time was the graphical presentation of the information in color diagrams and chart visualization. In 1896, the Statistical Bureau published an ‘Album of Graphical Statistics. The Situation of the population according to the census from 1st of January 1893’ (Statistical Bureau, State press 1896). The album included 36 figures and diagrams, and a French text in addition to the Bulgarian. During the world exhibition in Anvers in 1894, the album was awarded with an honorary diploma in the Section for Scientific Subjects and Cartography.

The structure and information included in the population censuses in Bulgaria improved and in 1892, questions about dwellings, stocks and agricultural households were

⁴A brief history of the National Statistical Institute of Bulgaria can be found on its homepage: <http://www.nsi.bg/AboutNSI/AboutNSI.htm>.

added to the census program. Moreover, the Department for Statistics developed a special 'Law for Conducting Censuses', which passed through the National Parliament in 1897. This law focused on the organization and structure of the censuses, and it specified the rules for conducting censuses, and for the methodology and data collection process, data protection issues and rules for data use. In summary, this law built up the regulations for conducting all Bulgarian censuses up to the beginning of World War II.

The first population census after World War II was conducted in 1946. This census was carried out in the tradition of the earlier censuses of the pre-war period. In the next census conducted in 1956, a new procedure for obtaining preliminary results from the census was applied: a 5 per cent sample was drawn from the census; six months after the census the Statistical Institute published preliminary results from the 1956 census. The procedure to publish preliminary results from the census became routine in the following years and is still applied today. However, until 2001 when the sample size drawn from the census was increased to 2.5 per cent, the preliminary results from the censuses after 1956 were published on the basis of a 2 per cent sample.

The 1965 census (as well as the following two censuses in 1975 and 1985) was innovative in the sense that it was complemented and accompanied by specialized representative surveys on internal migration, fertility and reproduction, a survey on the health status of the Bulgarian population and a labor activity survey. This procedure became the usual practice and since 1965, various specialized representative surveys were taken parallel to the census enumeration.

The main effort of the census in 1992 was to harmonize the enumeration program in Bulgaria with the census programs in other European countries and in particular with the recommendations of the International Statistical Institute. This census was the first one taken after the transition towards a market economy started in Bulgaria, and it documents, for example, the socioeconomic characteristics of the Bulgarian population in the beginning of this profound transition. Moreover, the census is part of the 'National Demographic Database' and was used for performing the first link between mortality and census records.⁵

4.1 The 2001 Population Census

The last Bulgarian census on population, housing and agricultural farms was conducted on March 1, 2001, and incorporated the recommendations of EUROSTAT, UN and other international research organizations. All data collected in the census refer to this reference day of the census (March 1st 2001, 00:00 AM). The organization and implementation of the most recent census program was carried out by the National Statistical Institute, which was responsible for the methodological preparation of the census as well as the data processing and publication of the results. The methodology implemented in the most

⁵Information about the exact day and year of all Bulgarian population censuses can be found on the homepage of the National Statistical Institute at <http://www.nsi.bg/Census/SrTables.htm>. Moreover, the tables presented on this homepage give the total number of the Bulgarian population for each census separately by sex and urban rural type of residence.

recent Bulgarian census was developed in cooperation with the Statistical Commission of the United Nations, the European Economic Commission and EUROSTAT. The National Statistical Institute started with the preparations of this census in September 1999 and employed 45,437 interviewers and supervisors during the census.

In the section below, we briefly describe and present the main targets and units of census enumeration in Bulgaria in 2001. In addition, we give some definitions and classifications that in our opinion are important to understand the questions and information collected in this census.

4.1.1 Targets and units of the census enumeration

The following individuals were enumerated in the census on March 1, 2001:

1. All Bulgarian citizens who live permanently in Bulgaria regardless of whether they were in the place of their permanent residence or in a different place on the reference day March 1, 2001;
2. All Bulgarian citizens who left Bulgaria after February 28th, 2000, as well as family members who accompanied them (by proxy interviews);
3. All Bulgarian citizens with a permanent residence in Bulgaria, who were officially representing Bulgaria, national institutions and other national organizations abroad on the reference day, as well as their family members;
4. Foreign citizens who have received a permanent residence in Bulgaria before March 1, 2000;
5. Foreign citizens and their family members who were granted political asylum and have received an official status as refugees in Bulgaria before March 1, 2000;
6. Bulgarian citizens and those with a dual citizenship who lived permanently abroad but were in Bulgaria on the reference day;
7. Foreign citizens who received an official residence permission after February 28, 2000;
8. Foreign citizens and their family members who had an open legal procedure for granting political asylum and status as a refugee according to the national and international law after February 28, 2000.

In addition, the housing census obtained information on:

1. Buildings including those which were not used for living purposes on the reference day of the census;
2. Inhabited and uninhabitable dwellings, and dwellings used for non-housing purposes;
3. Agricultural farms owned by the persons mentioned in point 1 and 2 above.

4.1.2 Census definitions and classifications

1. Population observed in the census:

Permanent Population – includes all persons who live permanently in a given settlement, irrespective of whether they are present on its territory on the reference day of the census, or have temporarily left the settlement. The permanent place of residence is defined as the settlement in which the person usually lives.

Present population – includes all persons who are present in a given settlement on the reference day of the census, regardless of whether they live there permanently or are temporarily present for any reason.

Temporarily absent persons – are those who live permanently in a given settlement, but at the reference day of the census were absent from the dwelling because of business, vacation, or other trips.

Temporarily present persons – those who live permanently in another settlement, but at the reference day of the census were in the respective settlement under census enumeration because of business, vacation, or other trips.

2. Households and families (households, nuclear families, position in the family and the household):

The basic unit of observation in the census is the *household*. A household can be comprised of one or more people. A *one-person household* consists of one person who lives in a self-contained dwelling or in a room or section of a dwelling. This person lives on a self-dependent budget for food and other expenses. Persons who live permanently in worker or student hostels are also considered a household.

A *household consisting of two or more people* combines people who live together in a dwelling or part of a dwelling, have a common budget for food and other daily necessities, and eat meals together regardless of the fact that some of them may not be related to each other.

A *collective household* consists of a group of people who live permanently in a collective, have a common budget, eat meals together, and are committed to a common living regime (i.e., people living in monasteries, houses for disabled people, orphanages, etc.).

According to the census definition, unmarried children, irrespective of their age, who live with their parents are considered members of their parents' household. Divorced or widowed persons who live with their parents are considered members of their parents' household only if the situation meets the criteria for a household formation.

A common budget exists only if one or more members of the household use resources of the other members. Persons who do not share a common budget and do not share

common meals do not form a household, regardless of the fact that they may live in the same dwelling. Separated spouses whose marriage is not legally dissolved do not form a household, even if they continue to live together. However, if two or more families living in one dwelling meet the criteria for building a household, i.e. they have a common budget and eat meals together, then they are counted as one household.

A *head of the household* is considered the person who provides the main resources for living, or the one who has been accepted as a household head by the other household members (most often by the right of seniority). Usually, the head of the household is also the head of the family. If a husband or wife lives permanently in another settlement, then he or she cannot be head of a household. Kinship between the members of a household is not a requirement for forming a household.

A *family* may be composed of:

- (a) spouses, independently of whether they are legally married or live in a cohabitation, without children, or
- (b) spouses with one or more never married children regardless of the children's age, or
- (c) a single parent with one or more never married children regardless of the children's age. The definition of family accepted in Bulgaria meets the requirements of the United Nations for observing nuclear families. According to this definition, there is a distinction between a biological family and the nuclear family, which consists only of parents and children who live with their parents at the moment of the census and have never been married. There are no limitations concerning the age of the children or their economic activity at the reference day of the census.

The 2001 census in Bulgaria differentiates between the *de jura* and *de facto* marital status. According to the census definitions, the legal marital status (*de jura*) is defined by an official certificate for civil marriage, divorce, or a death certificate for one of the spouses. The *de facto* marital status differentiates between living in a marriage, living without a marriage or not in a marriage, irrespective of whether this marital status is legally formalized or not.

According to the census methodology, a household can have the following composition:

- (a) Head of the household – the person determined by the household as a head;
- (b) spouse of the household head, independently of whether legally married or not;
- (c) son/daughter of the household head permanently living in the household, including stepchildren;

- (d) son/daughter-in-law – son/daughter-in-law who are respectively husband/wife of the household head’s children and who live permanently in the household;
- (e) grandchildren of the household head who live permanently in the household;
- (f) parents of the household head or his/her spouse;
- (g) grandmother/grandfather of the household head or his/her spouse;
- (h) other relatives – persons who are in a kinship with household head or his/her spouse and who live permanently in the household;
- (i) other persons – this category includes all persons who are not in a kinship with the household head. This category also includes all members of collective households (boarding houses, hostels, homes for elderly or ill people, infant or orphan homes, monasteries, etc.).

According to the census, the family composition is defined as following:

- (a) Family head is the person determined by the family as a head;
- (b) Spouse of the family head, regardless of whether they are legally married or not;
- (c) Children – never married children, including stepchildren, who live permanently with their parents.

3. Ethnic affiliation:

The 2001 population census asks three questions related to the ethnic affiliation: ethnic group, mother tongue and religion of the persons who are subject to a census enumeration. *Ethnic group* is defined as a community of persons related to each other by their origin and language, and who are close to each other because of the customs of their life and culture. The census differentiates between three major ethnic groups in Bulgaria – Bulgarian, Turk, and Roma, and the question includes a category for others. *Mother tongue* is defined as the language a person speaks best and which is usually used for communication with the members of his family or household. The census differentiates between the three major languages spoken in Bulgaria (Bulgarian, Turkish and Roma), and gives the option for other languages. *Religion* is defined as the historically determined belonging of a person (or of his parents or grandparents) to a given group with certain religious views. The main categories given in the census form are Eastern Orthodox, Catholic, Protestant, Muslim Sunnite, Muslim Shiite, and others.

Answering the questions about ethnicity, mother tongue and religion is voluntary. If the person decides not to respond to these questions, the census enumerator marks the answers ‘not stated’. The ethnic group, mother tongue and religion of children are determined by their parents. If the parents belong to different ethnic groups, religions or have a different mother tongue, then they decide to which one their

children belong. The head of the household determines the answer to these three questions for deaf-and dumb or mentally disabled individuals.

4. **Employment, unemployment and sources of means of living:**

The 2001 census includes extensive information on employment, unemployment and sources of income of the population. One can differentiate between the economically active and economically inactive population at census. The *economically active population* includes all individuals who are above age 15 and are currently employed or are searching for a job. That is, this category includes employed as well as unemployed people. In contrast, the *economically inactive population* comprises all individuals who at the reference day of the census were not employed, did not receive any income from salaries or wages, and were not looking for a job.

According to the census definitions, the term '*employed*' refers to those persons who at the reference day of the census (March 1, 2001) were over 15 and worked and received a salary regardless of whether they were employed full-time or part-time. The category '*employed*' also includes those individuals who at the reference day of the census were absent from work due to different forms of leave, or were on a business trip, took qualification or retraining courses, or were absent for other reasons that do not cause discontinuation of the labor contract.

In contrast to the above definition, *unemployed* are persons above age 15 who at the reference day of the census (March 1, 2001) had no paid work, but wished to work and were capable of work and were actively searching for a job. Persons who have completed their education or have interrupted it and are not working but are searching for a job are also counted as unemployed regardless of whether they are officially registered as unemployed in the 'Unemployment Bureaus'. The category of '*unemployed people*' does not include people who receive a pension because of retirement or another kind of an individual pension.

The census distinguishes between the following types of employment:

Employers – persons who manage their own business (individually or jointly with other people) with employees;

Self employed – persons who manage or work in their own business (individually or with other people) without employees;

Employees – persons who work for an employer (with or without a working contract) and receive a compensation for their work in the form of wages, salaries or payment in kind;

Family workers – persons working in a family business or business of other relatives without any kind of payment;

Members of agricultural co-operatives – persons who work not only for income in kind, but receive also a salary/wage;

Others – persons whose type of employment cannot be defined by one of the above categories.

Each person participating in the census was obliged to state at least one source of means of living. In case persons had several sources of income, they had to state all of them. The 2001 census defines the following categories of sources of means of living:

Main job – the work a person does for a salary/wage or income in kind. In case a person has two or more jobs, the job in which he or she spends most of the working time during the day is defined as the main job;

Secondary job – the overtime work (different from the main job) done by a person for any kind of income;

Unemployment benefits – benefits received from the state or other organization if the person is officially registered as unemployed;

Pension – the census differentiates between the two types of pension – a pension that an individual receives because of retirement, and a pension for disability;

Maintenance from an employed person – this category includes all people who receive means of living from employed people;

Maintenance from an unemployed person – this category includes all people who receive means of living from unemployed people;

Maintenance from an inactive person – this category includes all people who receive means of living from people who do not work, but are not officially registered as unemployed;

Social maintenance – this category includes all people who receive means of living from the state or public organizations such as social benefits, unemployment benefits etc.;

Scholarships;

Means of living from interest rates, savings, loans, dividends from shares ,etc.;

Sources of means of living from an individual agricultural farms;

Others – sources of means of living which cannot be traced back to one of the above categories.

5. Migration:

The 2001 census collected also data on internal and out-migration of the population.⁶ Migration is defined as the change of the settlement of the permanent residence of

⁶Besides that, the census was accompanied by an extensive survey on migration, which we will not describe in this paper.

a person. The question asked in the census is whether a person has migrated in the period 1993–2001 (the period between the previous (1992) and current Bulgarian census). If a person has migrated within this period, then the census records information on the municipality, settlement, and possibly the country of a previous residence as well as the year when the person last migrated. The previous residence of children is recorded according to the previous (permanent) residence of the mother. If a person has migrated more than once during the period of observation, then the census records only the year and the destination of the last migration. Thus, according to this question on migration, the 2001 census gives information on the total number of migrated people and not on the total number of migrations within the period of observation.

Short stays (less than a year) of persons abroad or in another settlements due to business, vacation, visits, holidays, etc. are not considered migrations. Persons who are on a regular military service, in prisons, or in specialized boarding-schools or boarding-houses are also not considered migrants. In addition, changes of permanent residence due to administrative reasons because of merging or splitting, or creating new settlements, as well as ‘moving’ settlements from one to another municipality, are also not registered as migration.

In addition to the topics discussed in this part of the paper, the last Bulgarian census asks detailed questions on the type of dwellings, durable consumer goods possessed by the household, use of public and private transportation etc. However, the discussion of this information is not within the scope of this paper.

4.2 Post-enumeration Surveys on Coverage and Accuracy of Registration in the Censuses

Immediately after the census is completed, the National Statistical Institute carries out post-enumeration surveys in order to obtain information for the coverage and accuracy of registration of the information in the census. The first post-enumeration survey for assessing the quality of the census in the history of the Bulgarian census programs was taken after the Bulgarian census on December 1, 1965. Ever since, these post-enumeration surveys have followed each Bulgarian census. The main purpose for conducting the post-enumeration surveys is to collect statistical information about the national coverage of the census itself and possible content errors embodied in the census data. However, the information obtained from the post-enumeration surveys is not used for correction of the collected census data but only for making the users aware of possible sources for errors in the census data.

The post-enumeration surveys have a twofold purpose:

1. First, the intention of these studies is to scrutinize to what extent the census covers the entire population that is subject to a census enumeration. In this context, the

absolute and relative share of people who have been missed in the census or have been counted more than once is calculated. In addition, the post-enumeration surveys are used to calculate the errors made when people were assigned to the wrong type (category) of population at the census day. For example, these are individuals who were supposed to be a subject of census enumeration, but have been incorrectly classified as a permanent versus present population, or they have been omitted from the census. Alternatively, these are individuals who were not supposed to be included in the census population, but were erroneously counted in the permanent or respectively present population. Subjects of the post-enumeration survey for the analysis of the coverage error in the census are all individuals who are permanent or present residents of the dwelling, or are temporarily absent or temporarily present in the dwelling considered in the post-enumeration survey, regardless of their location on the census day itself.

2. The second purpose of the post-enumeration surveys is to investigate the errors related to the erroneous recording of personal or household characteristics in the census. Subjects of the post-enumeration surveys for the analysis of the accuracy of information registration in the census are all individuals who were counted as permanent population in the dwelling on the census day (e.g. permanent and temporarily absent population), or dwellings that are registered as a subject of census enumeration.

The most recent Bulgarian census from March 1, 2001 was also followed up by two post-enumeration surveys in order to investigate the coverage and content errors embodied in the collected data from the census enumeration. The post-enumeration survey for the analysis of coverage errors included 9,645 households with 26,722 persons. Its purpose was to estimate the errors due to counting some people twice or not counting them at all, or including some people in the wrong type of population. The post-enumeration for the estimation of the content errors in the 2001 census was based on 3,162 households with a total of 8,640 individuals. This survey covered the following personal and household characteristics:

1. Personal characteristics: age, de jura and de facto marital status, education, internal migration;
2. Dwelling characteristics: living space, subsidiary area of the dwelling;

The post-enumeration surveys for the coverage and accuracy of information registration of the censuses in Bulgaria are carried out on the basis of a two-level stratified sample, in which the first level is residential areas, and the second level is households. The sample is formed by place of residence: district, city/town, village. The type and size of the sample are defined so that they ensure the accuracy of the estimates on the basis of the observed variables in the sample, i.e. the relative standard errors are less than or equal to 20 per cent with a confidence interval of 95 per cent.

In summary, the concept and methodology of the post-enumeration surveys for the coverage errors and accuracy of registration taken in Bulgaria after each population census differ from the post-enumeration surveys implemented in the United States, for example (see for example Brown et al. 1999). In this latter approach one actually conducts an independent enumeration, using different lists and maps with sampling units and different personnel. In this way, it is actually possible to detect a dwelling unit that is missed in the census itself. However, the methodology that is applied in Bulgaria does not allow this, and if a dwelling unit is missed in the initial lists and maps of dwellings, it cannot be re-captured in the post-enumeration survey.

5 Specialized Demographic Studies

In order to understand the demographic processes in the country more profoundly, various representative demographic studies are carried out in Bulgaria. The most important of these studies are those that are supplementary to the censuses. The first demographic study that has been conducted parallel to a census focused on internal migration and was taken in 1965. Two demographic studies accompanied the census in 1975. One study focused again on internal migration, while the second study was a part of the UN program and focused on fertility.

In the following decade the program of these specialized demographic studies in Bulgaria was extended, and during the census in 1985, four demographic studies were taken parallel to the census: a migration study, fertility and reproduction study, labor and labor behavior of the population, and a study on the health status and health behavior of the Bulgarian population.

The transition period after 1989 is associated with a large reorganization of the statistical system in the country as well as with a serious limitation of the available finance resources. Thus, in 1992, the year of the next population census in Bulgaria, only two representative surveys had been carried out: a migration survey that focuses for the first time on out-migration, and labor force survey. Both studies, however, were taken separately from the census.

During the last Bulgarian census in March 2001, several specialized studies were included in the census program and were conducted parallel to the census enumeration. The first study focused on fertility⁷, while the second one was the Health Interview Survey. In addition, a study of the internal and out-migration of the Bulgarian population was carried out. Moreover, a part of the census program is also a specialized study on ‘Time Use’, which started on October 15, 2001 and continues until October 14, 2002.

The representative demographic studies are based on a permanent master sample of households in the country. This sample is maintained by the National Statistical Institute.

⁷The Max Planck Institute for Demographic Research in Rostock, and in particular researchers associated with ‘Division2’ headed by Professor Jan M. Hoem advised the questionnaire and the implementation of this study.

The sample procedure of the households is based on the data from the censuses and includes information on the number and composition of households, the demographic characteristics of the household members, and the address. The master sample includes 10 per cent of all enumeration districts in the country, and it is stratified by districts, municipalities and type of residential areas (e.g., cities, towns, villages). Parts of the master sample are updated regularly, depending on the frequency of its use. Structural changes in the demographic characteristics of the Bulgarian population are considered when the master sample is modified and updated.

The first master sample was established during the census in 1985. In 1992, the master sample was considerably modified while considering the enumeration results from the census. The master sample supplies the basis for various surveys performed by the National Statistical Institute, e.g. the regular household budget surveys (surveys on household income and expenditures), labor and unemployment surveys, etc.

The master sample also provides the basis for several other surveys conducted upon a requirement from other Bulgarian institutions or international organizations. For instance, in 1992 and 1996 the International Organization for Migration conducted studies on potential outside migration in Bulgaria. On the initiative and with the financial support of the UN statistical Office, in 1995 was conducted a social survey investigating ‘Men and women during the transition’. In the beginning of 2000, the Bulgarian Reproductive Health Survey was carried out. The survey was financed by a research grant from the U.S. National Institute of Child Health and Human Development and the University of South Carolina.

6 Data Protection and Data Confidentiality of Individual-Level Data in Bulgaria

The initial demographic data collected by the National Statistical Institute are on individual levels and contain extensive and quite detailed personal information. In order to ensure and guarantee the confidentiality of individual-level data, the National Statistical System applies various measures for data protection. We can distinguish among normative, organizational and technological measures.

The normative protection of individual-level data is ensured by virtue of the existing law that regulates the statistical informational system in Bulgaria, the process of data collection and data use, as well as the internal order of the National Statistical Institute, which acts as the institutional organization for collecting statistical information in Bulgaria.

According to the law for statistics passed in 1999, the individual-level data collected in surveys and censuses can be used only for statistical purposes. The statistical institutions and their representatives are prohibited from divulging individual-level information or data that might reveal the identity of a person. Employees working with individual-level information are obliged to keep secret all personal information and are obliged to sign a

statutory declaration for keeping official secrets. This obligation of the employees remains after the employees have quitted their job in the statistical system or the survey ends.

The law assigns administrative penalties and fines for employees who violate their official commitments for the confidentiality and protection of individual-level data. In some cases, the infringement can count as a crime and has other consequences.

The law for statistics regulates that the statistical organizations are responsible for the security and protection of individual-level information and are obliged to prevent any abuse of this personal information. Moreover, it is the responsibility of these organizations to provide the respective organizational and technical measures for ensuring the required data protection.

The 2001 population census is legalized by a separate 'Law for census on population, housing and agricultural farms' that regulates the respective preparation, organization and implementation of the 2001 census in Bulgaria. The law stresses in particular the topic of protection and confidentiality of individual-level information that is collected in the census. This individual-level data obtained from the census enumeration is protected in fact by the separate law for the census as well as by the law for statistics. Data collected during the census cannot be used by the executive authorities as an incriminating evidence against individuals.

During censuses, the National Statistical Institute employs a large number of people to collect and process data. Similar to the law for statistics, the law for the 2001 census defines strictly their responsibilities. Accordingly, the staff responsible for the data collection, data control and data processing during a census is prohibited from divulging and passing on individual-level information obtained from the census forms. Similar to other employees in the statistical system who work with individual-level data, they are also obliged to sign statutory declarations, and the respective obligations also continue after the end of their working contracts.

According to the law for the implementation of the last Bulgarian census, the institutions responsible for the census implementation are required to take all the organizational and technical measures for the protection of the collected personal information. Moreover, an article from the law says that all original census forms must be destroyed following special procedures three years after the census.

The National Statistical Institute itself implements special measures for guarantying the confidentiality of individual-level information. For instance, there is a special internal regulation system for use and supply of statistical information to outside parties. Moreover, only authorized staff has access to the original individual-level information stored in the database of the Institute.

Third parties, such as various national and international, governmental and private institutions, scientific organizations and others, can apply to use data from the demographic database at the National Statistical Institute. This procedure is regulated by the instructions for informational services at the NSI. According to regulations in the instructions, the potential users of statistical data need to submit an application in which they describe

exactly the content and structure of the required data. After the application is approved by the president of the National Statistical Institute or his/her authorized representatives, the respective data is prepared by the Department for Demographic Statistics and supplied to the users under consideration of the data confidentiality issues. This usually implies that the data is supplied to the users in an aggregate form so that no individual included in the dataset can be identified.

The structure of the demographic database at the National Statistical Institute and the availability of the personal identification number (PID) allow for linking various demographic events and demographic information. The first time this possibility was explored was for the supplement of the dataset used by Iliana Kohler for her dissertation research on mortality in Bulgaria during the transition period (Kohler 2001). The actual linkage of the data must take place at the National Statistical Institute by the respective experts who have an authorized access to the individual-level data. Moreover, employees of the National Statistical Institute decide which various indicators can be in the dataset and how the grouping of individual characteristics should be performed so that no individual can be identified. Furthermore, personal information such as the personal identification number, address and other information are not given to outside users of the demographic data. If users require individual-level survey data that includes individual records, the users sign a contract with the National Statistical Institute, which defines the conditions and restrictions for using the survey data (e.g., the contract defines the purpose of the study, publication of data, access to the data and confidentiality issues).

Acknowledgments: We thank the Max Planck Institute for Demographic Research in Rostock and the National Statistical Institute of Bulgaria for their support in making this collaboration possible. We are particularly grateful to the Max Planck Institute for providing financial support for this research report. We are also grateful to the colleagues from the Department of Demographic Statistics at the National Statistical Institute of Bulgaria, who are developing the methodology for collecting demographic data in Bulgaria and who are processing and developing the demographic database in Bulgaria. We thank Ivan Balev, Elka Jakimova, Mariana Mourgova, Dimiter Philipov, Ward W. Kingkade and Hans-Peter Kohler for many helpful comments and suggestions.

References

- Andersen, T. F., M. Madsen, J. Jørgensen, Mellekjær, and J. H. Olsen (1999). The Danish National Hospital Register. *Danish Medical Bulletin* 46(3), 263–268.
- Brown, L. D., M. L. Eaton, D. A. Freedman, S. P. Klein, R. A. Olshen, K. W. Wachter, M. T. Wells, and D. Ylvisaker (1999). Statistical controversies in census 2000. Technical Report 537, Department of Statistics, U.C. Berkeley.

- Eurostat/Statistics Denmark (1995). *Statistics on Persons in Denmark: A Register-Based Statistical System*. Copenhagen: Statistics Denmark.
- Frank, L. (2000). Epidemiology: When an entire country is a cohort. *Science* 287, 2398–2399.
- Knudsen, L. B. and M. Murphy (1999). Registers as data sources in studies of reproductive behaviour. Danish Center for Demographic Research Research Report #12.
- Kohler, I. (2001). *Adult and Old-Age Mortality Dynamics in Bulgaria and Russia*. Dissertation, Faculty of Social Sciences, University of Southern Denmark, Odense.
- National Statistical Institute (2001). *120 godini bulgarska statistika. 1880–2000. (120 Years Bulgarian Statistics)*. Sofia.
- National Statistical Institute of Bulgaria (1999). *Statistical Yearbook 1999*. Sofia.
- National Statistical Institute of Bulgaria (2000). *Statistical Yearbook 2000*. Sofia.
- Petersen, J. K. (2000). The Danish Demographic Database – longitudinal data for advanced demographic methods. Research Report # 15, Danish Center for Demographic Research, Odense.
- Statistical Bureau, State press (1896). *Album ot graphitcheska statistika. Sastojanieto na naselenieto spored prebrojavaneto na 1.01.1893g. (Album of Graphical Statistics. The Situation of the Population According to the Census from 1st of January 1893)*. Sofia.
- Statistiska centralbyrån (2001). *SCB-Data för Forskare. Individbaserade register med innehållsbeskrivning*. Statistics Sweden.