Action Plan to improve and enhance the migratory data production system and data sources in Hungary

Ádám Dickmann, Éva Gárdos, László Kajdi, Marcell Kovács, Erzsébet Eperjesi Lindnerné, Attila Melegh, Annamária Sárosi, Béla Soltész, Rita Váradi
Hungarian Central Statistical Office

Zsuzsa Blaskó, Irén Gödri
Demographic Research Institute at the HCSO

Katalin Füzér
Municipality of Pécs

Endre Sik
TÁRKI Zrt.

2014

Jointly for our common future

Information for reproducing excerpts from this report can be found at www.seemig.eu. Inquiries can also be directed to: Demographic Research Institute at the Hungarian Central Statistical Office H-1024 Budapest, Buday László u. 1-3. or by contacting dri-seemig@demografia.hu.

Suggested citation: HCSO, HDRI, Pécs, TÁRKI (2014): *Action Plan to improve and enhance the migratory data production system and data sources in Hungary.* Action plan developed within the project 'SEEMIG Managing Migration and Its Effects – Transnational Actions Towards Evidence Based Strategies'.
http://www.seemig.eu/downloads/outputs/SEEMIGActionPlanHungary.pdf

List of acronyms, abbreviations and their English translation

Acronym/ Abbreviation	English translation	Endonym
COAEPS	Central Office for Administrative and Electronic Public Services	Közigazgatási és Elektronikus Közszolgáltatások Központi Hivatala (KEK KH)
DRI	Demographic Research Institute	Népességtudományi Kutatóintézet (NKI)
EEA	European Economic Area	Európai Gazdasági Térség (EGT)
EU	European Union	Európai Unió
HCSO	Hungarian Central Statistical Office	Központi Statisztikai Hivatal (KSH)
ILO	International Labour Organization	Nemzetközi Munkaügyi Szervezet
LFS	Labour Force Survey	Munkaerő-felmérés (MEF)
NHIFA	National Health Insurance Fund Administration	Országos Egészségbiztosítási Pénztár (OEP)
OIN	Office of Immigration and Nationality	Bevándorlási és Állampolgársági Hivatal (BÁH)

Table 1. Overview table of the activities related to the WP4 Transnational Database

Activities	Level of intervention	Relevant stakeholder	Legal basis	Previous attempt for this type of intervention? If yes, why did not happen?	Possible difficulties	Rank of suggestion
Possible construction of a national developmental database on migration, labour market and human capital	national	HCSO, SEEMIG partners, Statistical Offices of the partner countries	TBA	No such attempt	Funding, data insecurity	
Possible construction of local databases on migration, labour market and human capital	local					
Maintenance of SEEMIG transnational databases beyond SEEMIG's lifespan	Institutional					

Table 2. Overview table of the suggestions to improve migratory data productions systems and major data sources

Data source type	Activities	Level of intervention	Relevant stakeholder	Legal basis	Previous attempt for this type of intervention? If yes, why did not happen?	Possible difficulties	Rank of suggestion
Administrative data sources	• Legal authorization of the access of the relevant statistical offices/ institution to individual data - changing of legal environment, modification of Act CXII of 2011 on data protection so that data owners have the right to share wider scope of data among themselves - modification of Act on Statistics No. XLVI of 1993 in order that statistical institution can have a general right to access non-statistical data to the extent necessary for the development, production and dissemination of national and European statistics.	national government	Data owner institutions involved in the integration of data sources	Act CXII of 2011 Act No. XLVI of 1993	Act LXIII of 1992 was modified in 2004 to comply the respective EU directive. No. Preparation of changing the act is running	opposing attitude of data owner institutions; communication towards the general public	
	•Steps towards creating better integrated data sources (integrated data sets, parallel use of several data sources) - the harmonization of time periods of European regulations with definitions applied in registration certificates	European	Eurostat	EU Regulation (EC) No 862/2007 Act I of 2007			
	• Steps towards the harmonization of	national	Data owner	Act XX of 1996	There is a project currently running		

Data source type	Activities	Level of intervention	Relevant stakeholder	Legal basis	Previous attempt for this type of intervention? If yes, why did not happen?	Possible difficulties	Rank of suggestion
	definitions in different migration data sets	data owner institutions	institutions involved in the integration of data sources		with this aim (EKOP 2.3.8)		
	• COAEPS should receive the data of third-country nationals having residence permit in Hungary	national data owner institutions	COAEPS, OIN	Act LXVI of 1992 Act II of 2007			
	• NHIFA should provide personal data to HCSO	national data owner institutions	NHIFA, HCSO	OSAP Act LXXXIII of 1997		opposing attitude of data owner institution; communication towards the general public	
	• Steps towards an integrated statistical database at the HCSO - Improvement of the consistency of data on asylum seekers, i.e. HCSO receives the Eurostat data-provision of OIN on asylum-seekers and persons granted international protection status	national	HCSO	Act XLVI of 1993	the project entitled 'Creating the basis for migrant-specific statistical data collections'	receiving personal data from the NHIFA ensuring human resources	
	• Dedicating extra financial resources for extra labor costs of maintaining national database on migration and development additional tasks	national government	data owner institutions			lack of financial resources	

Data source type	Activities	Level of intervention	Relevant stakeholder	Legal basis	Previous attempt for this type of intervention? If yes, why did not happen?	Possible difficulties	Rank of suggestion
	• Collecting data on previous / next country at the COAEPS	national data owner institutions	COAEPS, OIN	Act LXVI of 1992		opposing attitude of data owner institutions	
	• Steps towards the introduction of PINs of international migration statistics	national	HCSO, data owner institutions	Act XLVI of 1993 On Statistics		Opposition of data owners	

	•Steps toward a better knowledge of data producers/users of relevant legal regulations with regard to emigration and immigration (e.g. repeated trainings)
	
	  − Organizing training programs for statisticians on the legal regulation
	
	  − Forums for all the relevant institutions
	
	  − Continuous cooperation between statisticians and data users
		national	HCSO, data owner institutions				

	• Steps toward improving data collection procedure, increase coverage and quality of data, updating the databases
	
	  − The most extended use of disposable resources
	
	  − Common meta data system of used data sources
		national	HCSO, data owner institutions				

Data source type	Activities	Level of intervention	Relevant stakeholder	Legal basis	Previous attempt for this type of intervention? If yes, why did not happen?	Possible difficulties	Rank of suggestion
	− Searching for new data collection methods to capture emigrants as completely as possible − Improvement of the quality and compiling quality reports of used data sources						
	• Steps toward the inclusion of migration-related questions into administrative and statistical data collections (e.g. country of birth, country of birth of parents etc.).	national	HCSO, data owner institutions	Acts regulating administrative registers		Opposition of data owners	

HCSO	• Steps toward making better estimates
	
	   − Establishment of working group
	
	   − Using mirror statistics and other estimations
	
	   − Methodological improvement
		national	HCSO			Lack of resources at the HCSO	

	• Steps toward the improvement of data production process between phases of data provision for authorities till the phase of data publication (e.g. providing help for migrants when submitting various applications) − Review of data collection procedure from the point of view institutional regulations	national	HCSO, data owner institutions	Acts regulating administrative registers		Lack of resources and opposition of data owners	

Data source type	Activities	Level of intervention	Relevant stakeholder	Legal basis	Previous attempt for this type of intervention? If yes, why did not happen?	Possible difficulties	Rank of suggestion
	and the needs and problems of data providers (i.e. migrants) − consultation among data owners and HCSO						
Labour Force Survey	• Steps toward increasing the availability/coverage of foreigners by mitigating the language barriers during the fieldwork - Action already started	National attempts but transnational harmonizatio n would be desirable.	HCSO, potentially other national Statistical Offices, Eurostat	OSAP 1539.	Low response rates experienced so far.	Self-administration might lead to lower data quality.	
	•Steps toward boosting the immigrant sub-sample and adjusting it to the territorial distribution of immigrants - Action already started	national	Financing body, HCSO	-	-	Low quality of administrative registers that could potentially serve as sampling frame.	
	• Steps toward the harmonization of the definition of household membership - Refinement and additional questions needed (see in point 3.2)	national	HCSO	-	-	intervention will affect time-series data	
	• Steps toward the standardization of concepts for identifying foreign-born	-	-	-	-	-	

Data source type	Activities	Level of intervention	Relevant stakeholder	Legal basis	Previous attempt for this type of intervention? If yes, why did not happen?	Possible difficulties	Rank of suggestion
	population - already standardized						
	• Steps toward the inclusion of supplementary questions regarding emigration or labour force out-migration - additional battery based on SEEMIG pilot experiences to be added periodically (minimum 2 years) - further questions refining current categories of labor market migration to be added to the core questionnaire, supplementing it	national and transnational national	HCSO, national statistical offices, Eurostat	-	SEEMIG LFS pilot study	avoiding low response rates harmonisation across countries	
	•Steps toward the harmonization and wider selection of non-response codes, in order to identify when the whole household went abroad. - national extension of the code-scheme needed, international harmonization recommended	national and transnational	HCSO, national statistical offices, Eurostat	HCSO's standard	-	-	
Migrant-specific surveys	•Introducing systematic, internationally standardized, survey-based migration-						

Data source type	Activities	Level of intervention	Relevant stakeholder	Legal basis	Previous attempt for this type of intervention? If yes, why did not happen?	Possible difficulties	Rank of suggestion
	related data collection be introduced - a specific module to be linked to LFS to assess emigration	see above	see above	see above	see above	see above	
	- harmonized immigrant survey in the main immigration countries	transnational	research institutes + external funding bodies	-	-	establishing international cooperation + getting funded	
	•Designing surveys at transnational, national and on a local level - longitudinal immigrant surveys on the national level	national	national research institutes + funding bodies		-	developing appropriate sampling frame, getting funded	
	- smaller-scale surveys on emigrants	national			a couple of small-case studies	developing appropriate sampling frame, getting funded	
	- surveys to be designed to understand the wider social context of migration processes	national			-	getting funded	
	- possibility to make surveys linking labour market, migratory and family formation strategies locally						
	- possibilities of applying the SEEMIG pilot methodology to local surveys					financing, involving methodological expertise on local	

Data source type	Activities	Level of intervention	Relevant stakeholder	Legal basis	Previous attempt for this type of intervention? If yes, why did not happen?	Possible difficulties	Rank of suggestion
						level	
	• Possibilities of planning modules for already existing international surveys (for instance related to outmigration) - new modules with migration focus to be added to international social attitude surveys e.g. ESS, ISSP which link labour market and migration issues	transnational	the transnational academic community involved in migration research		ESS module in 2002		
	• Steps towards setting up longitudinal migration relevant survey						
	Possibilities of conducting national longitudinal surveys based on local government questionnaires						
	OTHER: setting up household surveys to study the role of remittances.	national	national research institutes + funding bodies		-	developing appropriate research method + getting funded	
Census	•Steps towards the harmonization of questions referring to migration history	Not necessary, since it complies with EU Regulation (EC) No					

Data source type	Activities	Level of intervention	Relevant stakeholder	Legal basis	Previous attempt for this type of intervention? If yes, why did not happen?	Possible difficulties	Rank of suggestion
		763/2008.					
	•Steps towards the inclusion of supplementary questions regarding immigration (legal status of foreign nationals at the date of arrival and at the time of the census; date of acquiring citizenship; recording all citizenships(in the case of individuals with multiple citizenship; country of birth or former citizenship of parents)	EU level national level	EUROSTAT HCSO Government	EU census regulation national census legislation	The project Developing the system of immigration and integration statistics (EIA/2007/3.2.3.1) in 2009 already formulated some suggestions in this respect. Because the questionnaire was over-burdened already, only one question was included in the 2011 census.	too long questionnaire; migration is one of many core topics of the census; overburdening of data providers	
	• Enhancing the successful enumeration of foreigners living in Hungary	national level	HCSO	national census legislation	Steps were made in the 2011 census (see text)	Immigrants living in Hungary form a heterogeneous group, but are small in number and geographically scattered. the number of foreign-languages available in the online questionnaire is limited the difficulty of employing enumerators with foreign language competence	
	• Steps towards enhancing data on emigration and inclusion of supplementary questions regarding out-migration & return migration.	national level	HCSO government	national census legislation	Questions regarding emigration were first included in the 2011 census (for details see text). This, however, can only be used for estimation, so further improvement is needed.	increasing burden of data provision; data collection is difficult (or even impossible) if there is no one to provide data about the person living abroad;	

Data source type	Activities	Level of intervention	Relevant stakeholder	Legal basis	Previous attempt for this type of intervention? If yes, why did not happen?	Possible difficulties	Rank of suggestion
	Census Hub						
	• Preparations to introduce register-based census.	national level	HCSO government owner institutions of various registers	The establishment of the new legal environment is necessary.	Possibilities have been reviewed, but realization has been hindered by legal environment and the present state of registers.	lack of cooperation between various data owner institutions; improving the content of registers is a long-term and costly process; lack of a uniform personal identification number; difference between registered and actual place of residence; change of legal environment is necessary	
Other	TBA						
	Remittances						
	Development of methodology how to use big data	national level	HCSO			new type of data sources, no experiences with it	

1. INTRODUCTION

The Hungarian action plan has set up a complex strategy to improve data collection and data use on a local and national level. In line with the strategic aims of the Wiesbaden memorandum, adopted by the Director Generals of the National Statistical Institutes (DGINS) in 2011, the Hungarian SEEMIG team is firmly convinced that statistics on migration and related demographic, human capital and labour migration processes can only be improved if the intervention contains various targets and multiple levels of intervention. This is because statistics will remain highly problematic and there are no methods that can claim a very high level of reliability. The Hungarian SEEMIG team is also sure that even the best national complex strategies are not fruitful if the Hungarian data systems are not integrated at higher level into the transnational databases developed by the UN, OECD, Eurostat, World Bank. The Hungarian SEEMIG team also believes that the role subnational, regional and local strategies are going to gain importance and the linkages between national research institutes, public bodies (like statistical offices and local regional bodies) and researchers should be developed if reforms want to achieve the desired impact. The Hungarian SEEMIG team has been also trying to map related projects past, present and foreseen future in order to capitalize on these related efforts.

The Hungarian SEEMIG team has set the following comprehensive key targets to be achieved in the next five year period.

1. To improve migration related data in multiple sources (Labour Force Survey, Census, relevant registers) via introducing new, supplementary questions on migration history and via pushing for the introduction of questions on citizenship and country of birth in all those sources which have not completed this work. This work will be complemented with the required harmonization of definitions.

2. To identify and integrate new administrative data sources and to improve administrative data collection, data revision rules and procedures which would increase the quality and the coverage of the register data.

3. To make plans for linking register data via introducing a PIN and to work out a complex professional strategy to build up capacities being able to handle such integrated systems

4. To improve samples used in various data collections like LFS and other surveys via various techniques (including boosting samples, overcoming language barriers)

5. To plan surveys on transnational, national and local level taking into account the results of already completed surveys within and outside SEEMIG

6. To improve local data collection via planning new surveys and via utilizing the results of SEEMIG pilot techniques, on a local level.

7. To maintain SEEMIG database on migration, labour market and human capital and to assure its sustainability

8. To contribute to developing new techniques to estimate migration flows as a necessary and complementary technique to all other data collection methods.

9. It is also planned how "big data" techniques could be integrated.

2. POSSIBILITIES OF SETTING UP DEVELOPMENTAL DATABASES ON NATIONAL, LOCAL AND TRANSNATIONAL LEVEL

- *Possible construction of a national developmental database on migration, labour market and human capital*

The HCSO provides data for national users with several dissemination methods. Tables with fixed disaggregation can be downloaded for different periods of time (intra-annual, i.e. monthly/quarterly data, annual and time-series) and for different regional levels in the topic of demography and migration from the website of the Office. Another important tool is the dissemination database which offers dynamic tables for more experienced users in various breakdowns. The most important publication on migration data is the Demographic yearbook, which contains numerous very detailed tables on migration, asylum-seekers and naturalization as well.

In the framework of the SEEMIG project a comprehensive dataset were compiled including also Hungarian migration statistics for the period 2001-2011. The future maintenance of the dataset, the regular update and publication could be an important contribution to get a clearer picture on migration processes in Hungary. A future national national developmental database on migration, labour market and human capital could be founded on the SEEMIG transnational databases beyond SEEMIG's lifespan established.

- *Possible construction of local databases on migration, labour market and human capital*

See Chapter 4 for issues concerning local databases.

- *Maintenance of SEEMIG transnational databases beyond SEEMIG's lifespan*

In the framework of SEEMIG project an international migratory database was established with carefully selected indicators of the participating countries concerning migration, complemented by Eurostat and OECD data. Indicators included cover the domain of international migration and also related social and economic processes, mainly labour market and demography. Indicators provided by the partners were taken from a rich set of data sources including national LFSs, censuses, other national surveys and also administrative data. Collected indicators are now getting continuously updated on the SEEMIG website providing a uniquely selected set of time-series data from the eight SEEMIG countries covering the time period from 2001 to 2011.

To make maximum benefit out of this valuable project-output, it is very important that the database gets maintained even after the project's lifespan that is after 2014. As the greatest part of the

academic, as well as the financial investments has already been made, maintenance requires relatively moderate resources. The HCSO should take a leading role in this activity not only by coordinating the collection of the indicators from the partner countries on a yearly basis, but also by continuously considering possible areas of improvement. Besides reviewing the relevance of the indicators already covered, possibilities to expand the database by including further indicators should also be considered. Further areas of improvement include developing interactivity as well as new forms of visualisation possibilities (e.g. inclusion of maps and charts). The scope of relevant metadata should also be defined e.g. the specification of data sources, the inclusion of additional remarks provided by SEEMIG partners etc. Naturally, the HCSO should also take the responsibility for providing the Hungarian data and for collecting the relevant indicators from other international websites.

To carry out these tasks it is unavoidable to expand current capacities at the HCSO. Furthermore, maintenance of an efficient international partnership is needed with partners not only highly committed to the task but also with the necessary capacities and authorizations to access and process data needed for the purposes of the database. Coordination of such a partnership should also lie with HCSO.

3. NATIONAL LEVEL ACTIVITIES FOR MIGRATION RELATED DATA SYSTEM IMPROVEMENTS

3.1. Administrative data sources

•*Steps towards creating better integrated data sources (integrated data sets, parallel use of several data sources)*

Several previous projects have drawn the attention to the importance of the integration of different data sources. In the middle of the 1990s the HCSO initiated the cooperation of stakeholder organizations in producing migration statistics. As a first step the identifiable micro data files transferred from the Office of Immigration and Nationality (OIN) and the Central Office for Administrative and Electronic Public Services (COAEPS – Population register) enables the HCSO to link and validate the disposable data. European Union requirements were also an important driving force for this development. Based on the Government Decree No 2104/2004 an inter-ministerial commission on migration was organized, which clearly indicated the intentions of the government to establish an integrated migratory database. The Commission formulated recommendations on the importance of this task in line with the proposals of subsequent international migratory projects e.g. THESIM, MIGRINFO, PROMINSTAT. On national level two projects on the development of migration statistics in 2009 and 2012 also highlighted the fundamental priority of the integration of data sources.

The obstacles of database integration are multiple: different organizations collect data, data collections and definitions are not harmonized, and no common Personal Identification Numbers (PIN) is used that would help the HCSO to make data linkage. Thus although the need for integration was declared in several forums, the relevant stakeholders are not interested in development due to additional need for resources and the fear from emerging quality problems. The HCSO must undertake the data integration and maintains the database.

Data of third-country citizens currently handled by the OIN could be included in the database of the COAEPS to prepare and maintain a comprehensive population register, which contains all nationalized and foreign citizens with Hungarian residence. For this sake, the modification of Act LXVI of 1992 and Act I and II of 2007 might be necessary as well. The inclusion of asylum applications (OIN) and refugee stock data (COAEPS) in a common database could also improve the quality of data regarding asylum. There is a need for the improvement of the consistency of data on asylum seekers, i.e. the HCSO receives the Eurostat data-provision of OIN on asylum-seekers and persons granted international protection status.

The Hungarian government recently established a National Register as a web-based interface to contact with Hungarian nationals living abroad. Data are handled by the COAEPS, the further utilization in migration statistics should be examined.

The above mentioned issues underline the necessity of defining the scope of collected basic data, the definitions used, as well as the supplementary meta information, which would facilitate the use of data for statistical and research purpose. A standardized metadata system could be prepared, which supports the integration process.

The integration of existing administrative data sources can be promoted by Act CCXX of 2013 about the general rules of cooperation among registers of the government and local governments. The

long-term goal of this act is to ensure that organizations with public duties have registers built on generally prescribed criteria and contain standardized data. This law makes electronic data connection and data queries possible, except in case of data on taxation, data related to national security issues and qualified data. The act introduces no further access rights, only previously legitimized organizations can access data. Based on this act, however, the so called primary data sources have to be defined for frequently used data, primary meaning the most reliable data sources. Moreover, a register and system of meta-database will also be established. A shortcoming of the law is that it does not discuss data transfer needs that are beyond the building and maintenance of registers, although the majority of problems related to the statistical use of administrative data are such. However, discussions between the HCSO and the submitting body of the act have not concluded yet.

It is worth highlighting the difference between an integrated statistical database and an administrative database. The latter would imply that the participating organizations could access all the data of citizens, including e.g. health care or taxation, which is not a likely option due to data protection rules. The current 'Comprehensive client identification' project of the COAEPS therefore tries to solve the integration problem without creating an actual common database. Instead, it aims to create an interconnected register which contains only the key variables of the databases without personal data. Thus the product of the project is rather a standardized framework serving as a logical register instead of a central database. The importance of this project is that for the time being, there is no such standardized application which could support the handling of all administrative tasks related to a client in one system, e.g. there are separate databases for natural persons (population register, refugee register etc.) and non-natural persons (entrepreneurs, civil organizations, etc.). Therefore, a new IT framework is necessary to provide the clients of administrative services a standardized handling process. Besides the main priority i.e. improving client services, the project also supports making the authorized data linkage simple and fast. The modification of Act XX of 1996 might be necessary to entitle the HCSO to use the register of conjugate identifiers.

An integrated statistical database serves clearly just statistical purposes; hence the establishment is unequivocally the task of HCSO. The Statistical Office has access to the central immigration register, the population register and the taxation register and obtains personally identifiable data. However, data from the National Health Insurance Fund do not contain variables for personal identification, which would be important for an integrated database. To transfer names, addresses and dates of birth the modification of Act LXXXIII of 1997 on Health insurance is also necessary.

The establishment of an integrated migratory database is also part of the ongoing 'Social surveys on migrants' project, which will be closed on 31 June 2015.

• *Steps towards the introduction of PINs*

Following/Detecting the change in foreigners' legal status in registers and statistics has long been a problem. In order for faster and easier data connection, the introduction of a personal identification number seems to be necessary. There are various possibilities for this:

1. The re-introduction of the 11-digit personal identification number[1]. People's trust in the authorities and the favourable political climate is a precondition for this. At the moment, this does not seem to be the case. Even if only practical issues are considered, the use of the personal identification number would only increase general distrust.

2. Connecting various identification codes used by the relevant organizations, with necessary data protection protocol. The involvement of a methodological and a legal expert and ensuring the IT environment and procedures are necessary for developing this. It should be studied how the register of conjugate identifiers can contribute to such a system of the common identification for statistical use.

3. The introduction of a special personal statistical identification number (SIN) that all immigrants and their children born in Hungary would get. This SIN would only serve statistical and no police/legal purposes. A section on this could be inserted in Act II of 2007 or possibly Act CCXX of 2013, following the principle of interoperability. The SIN should be applied by the OIN and the COAEPS until the person becomes a Hungarian citizen.

4. ***If none of the above permanent solutions is feasible a method of temporary connection of personal, tax and social security data*** on a regular basis and only for statistical purposes should be developed, as the current methods of temporary connection are not suitable for setting up time series (anonymized records are stored which are re-labeled with a new hash code every 6 months).

•*Steps towards a better knowledge of data producers/users of relevant legal regulations with regard to emigration and immigration (e.g. repeated trainings)*

According to international experiences, the HCSO has to play an integrating role in providing information on international migration. The regular communication and discussion between data owners and the statistical office is indispensable. Regular professional trainings for the staff of the institutions related to migration, brainstorming on new methods with legal experts, researchers and statisticians are crucial. One option for this is the participation of staff involved in migrant-related issues in HCSO internal trainings, which are open for the members of Official Statistical Service. The elaboration of a migrant-specific optative module for the obligatory training course of public officials could also be useful.

Continuous cooperation between statisticians and data owners is a precondition for the effective use of administrative data for statistical purposes. In order to improve data quality, data owners should take into consideration the comments, experiences, needs as well as development and updating recommendations of statistical stakeholders. On the basis of these regular discussions, data-owners can assess the benefit of the proposals and are able to elaborate and execute development plans.

• *Steps towards the harmonization of definitions in different migration datasets*

---

[1] In 1975 the Ministry for Home Affairs started to provide people with PIN. In 1991 the Constitutional Court prohibited using the PIN generally. However the PIN is henceforward made up, it can exclusively be used by the Population Register. The Social Security and the Tax Authority have developed different identification numbers, respectively.

The applied definitions, classifications and metadata are often different in various migratory datasets as they are designed for different purposes. On international level the definitions of the Regulation (EC) No 862/2007 can serve as a guideline, but further clarifications at the national level are needed in order to use common terms. An important bias in data may be caused by the different thresholds for residence periods used by the European requirements on migration statistics (12 months) and by the registration certificates of EEA citizens (3 months). Although tests show that at the time being those who apply for these certificates stay in the country for more than a year, data quality could be improved if short term (i.e. max. 12 months) migrants could obtain different type of permits. The expiry of the documents after a certain period (e.g. 5 years) could also be a further important development, since migration databases could be updated.

Another practical example for the relevance of the harmonization is that the sub-registers of the OIN are not fully consistent, they cannot be interlinked. Data process is partly outsourced, which can be a major reason for the inconsistent data files sometimes transferred from the OIN to the HCSO. Continuously changing legal background is another reason that makes modifications in the structure of data files necessary. New variables should be introduced regarding comparability in time.

In previous studies on the development of migration statistics it was also revealed that for educational and occupational data used by OIN the classifications and categories used are too wide. In the framework of a standardized metadata system the ISCED and ISCO levels to be used should also be defined.

• Steps towards improving data collection procedure, increasing coverage and quality of data, updating the databases

Migration statistics have to follow the requirements of European statistics i.e. objectivity, reliability, impartiality, scientific independency, cost efficiency as well as the rules for data handling, data protection and confidentiality.

Quality reports must be compiled on the different data sources on migrant groups assessing their accuracy.

The first step is to extend the use of disposable resources. Registers contain the basic data needed to meet the requirements of national and international data collections. However, there are tasks to accomplish for further development:

• For the statistical use of data sources it is essential to harmonize the definitions used by data owner organizations and by the HCSO i.e. to create a standardized set of concepts, definitions and metadata for the different databases. The participation of the HCSO in the design and modification processes of the databases at relevant stakeholders could ensure the consideration of statistical aspects. An important development will be the interconnected register produced by the above mentioned COAEPS project, which will facilitate interoperability between data sets. During the preparatory process of Act CCXX of 2013, the HCSO was promised to be invited to participate in building the global meta-system of this register.

• Regular statistical surveys should be carried out on topics, which are not covered by registers. A good example for the need of developing surveys is the HCSO survey Nr. '1970' launched in 2003, designed for foreigners with residence permits (see SEEMIG Analysis of existing migratory

data production systems and major data sources in Hungary 2.2.5). The non-response rate is significantly high, mainly due to the circumstances of data entry and record as well as language barriers. Trainings for the responsible staff and translation of the questionnaire to foreign languages could improve response rates. As a first step a Hungarian – English bi-lingual questionnaire was introduced in 2013. The development of an online questionnaire and an online interface would also be useful for increasing coverage.

•   In line with the improvement of data content, data owner institutions should enhance the application of the principles of the European Statistics Code of Practice in order to provide better data quality.

•   During the data collection processes, variables which only serve statistical and not administrative purposes must be recorded in good quality. For instance in the data collections carried out by the OIN, important data on education and occupation are lost, since they are not used for administrative purposes. However, from a statistical point of view, these data would be essential.

•   The quality of data entry and data record should be improved – e.g. to avoid cases where key variables like citizenship are missing. A new data entry programme might be necessary to reach better quality with built-in quality checks (e.g. no proceeding allowed in case of missing answers) as well as the introduction of new questionnaires with closed questions.

•   Improving international cooperation is fundamental. Based on a previous experience between the Statistical Offices of Spain and Hungary, for instance, exchanging migratory datasets between EU member states could be fostered. This would help improving data quality and producing more accurate emigration numbers. The mutual informing on deceased citizens with other countries could also be useful from the data quality aspect. However, when attempting to exchange datasets among countries serious data protection issues generally emerge. As of now, the Spanish example, which makes the international transfer of migratory datasets possible, seems to be exceptional. Legal regulations do not authorize the NSIs to transfer such datasets in most countries.

•   Elaborating procedures to facilitate the control of foreign citizens residing in Hungary who have permits with no expiry date (e.g. registration certificate). Thus the under-estimation of emigration caused by the lack of deregistration and the disproportionate aggregation of foreigners could be adjusted by certain periods (2-5 years). The method can be a comprehensive full-scope or a sample survey. For the latter option estimation methods must be elaborated for settlement level.

•   The HCSO should obtain data from the National Health Insurance Fund on persons who did not pay social insurance for a longer period (e.g. one year). A statistical survey could be carried out on these persons, checking whether they are emigrants or have other reasons for not paying social insurance. Using data of tax office, too, those should be also included in this survey who do not pay tax but are at active ages, have working ability and are not in the unemployment register,

•   Administrative de-registration process of addresses (of possible emigrants) is done by the local notary, in the framework of a public administration process, when a resident does not reply to official written requests for a given period of time. However, national level authorities are not notified about this de-registration. It should be observed how information exchange could be resolved between the institutions.

• Steps towards the inclusion of migration-related questions into administrative and statistical data collections (e.g. country of birth, country of birth of parents, etc.)

Concerning migration-related questions, the introduction of the country of birth variable in 2008 (in case of death statistics 2010) was a huge step forward, since in certain cases it is hard to get a clear and realistic picture on migration processes based on data by citizenship. E.g. in the case of Hungary this issue is crucial since the introduction of simplified naturalization and the hundred thousands new Hungarian citizens residing abroad.

From the reference year 2014, according to the modification of the Act on Statistics, information on the mother's country of birth in the case of births, as well as the deaths of Hungarian citizens abroad are collected.

Another development could be the inclusion of data on previous and next country in the population register of COAEPS. This is currently carried out at the OIN, however, immigrants declare information about the country of residence that is planned to leave for when entering Hungary, which results in obviously inaccurate data. Data collection at the time of leaving the country would be more reliable.

It is worth mentioning that not only the scope of collected variables should be widened, but also the data record and the data process is inappropriate in case of certain variables, hence they cannot be used for further statistical analysis due to poor data quality.

Steps towards making better estimates

The number of emigrants is underestimated in most countries due to the fact that the administration procedures cannot follow a big part of people who leave the country. The consequent quality issues of different datasets (e.g. over- and under-coverage) require the elaboration of new estimation methods. In order to do that we suggest the following:

1. A working group must be established composed by methodology and migration experts.

2. For the estimations, foreign mirror statistics, estimations of NGOs and other relevant stakeholders as well as interviews with experts must be utilized besides official national data.

3. The possibility of surveying emigrants through Hungarian consulates abroad should be examined.

Steps towards the improvement of data production process between phases of data provision for authorities till the phase of data publication (e.g. providing help for migrants when submitting various applications)

• Introduction of administrative rules to support citizens and foreigners to follow legal instructions. Establishment of single channel or one-stop-shop client handling would be desirable, e.g. declaring deregistration only in one institution when leaving the country. Local support teams are also useful to solve language problems in administrative cases of migrants. Incentives can be used to increase the rate of deregulation among emigrants e.g. they would not lose allowances after leaving the country.

3.2. Labour Force Survey

The below plans and actions will be coordinated by the joint work of LFS and migration team of the Hungarian Central Statistical Office regularly reviewing the below process.

While in Hungary the number of immigrants is relatively low, and their territorial distribution is scattered there is a little chance for them of getting into the survey's sample. It is even more difficult if the household is not a Hungarian speaking one. So immigrants in the final sample of the HLFS do not represent the migrant population living in Hungary. To get a real picture on the situation of migrants from LFS a special sample is needed. Several current developments (in the frame of EIA/2013/2.6.1 project) are focusing at the improvement of better attainment of the immigrant population in Hungary.

Concerning emigration the core battery of LFS involves those persons whose workplace is abroad for one year but they are members of households living in the territory of the country. We could comply a special block of questions which helps better categorizing this special population and we could use as an ad hoc module from time to time (HCSO should find body or bodies financing this type of supplementary blocks.) This questions' block could also be used as a wave-specific block. Eurostat has a special methodological task force dealing with such proposals. HCSO also plans implementing a test on methodology of using wave approach[2].

Eurostat's reform plans for LFS are consulted with member states' statistical offices, which can propose the enhancement or reduction of the circle of currently used variables, and the topics and periodicity of the modules. The HCSO (Living conditions, Labour Market and Education Statistics Department) had a proposal to Eurostat in February 2014 in the framework of "Consultation rounds on the review of the EU LFS, spring 2014" to create a new ad hoc module for mapping migrants' demographic and labour market characteristics, or at least to enhance the present set of questions on the issue, in order to collect information not only inward, but outward migration as well.

•*Steps towards increasing the availability/coverage of foreigners by mitigating the language barriers during the fieldwork*

An initiative along these lines is already under way in Hungary. Although non-response due to language barriers is small (around 90 i.e. in a sample of 38 thousand addresses) an attempt has been made to eliminate this problem. Therefore the original LFS questionnaire was simplified to make it suitable for self-administering and then it was translated into 9 foreign languages. Interviewers were instructed to leave the translated copies in the foreign speaking households. The testing phase of this methodology has not been completed yet. Nevertheless we expect that the translation of the questionnaire is the most efficient approach to tackle the language problem. Since immigration in

---

[2] The *wave approach* is used for the *structural variables* which need be surveyed on a yearly basis, differently from *core variables* for which quarterly figures are required. In order to get for yearly structural variables the same reliability as for quarterly core variables, questions concerning the formers can be asked each quarter to a subsample of the full quarterly sample, e.g. the first wave.

Hungary is very low, employing foreign language interviewers would be too expensive. An experiment is planned using multi-mode data collection; web questionnaire will be tested as an option for non-Hungarian speaking population in the labour force survey.

•*Steps towards boosting the immigrant sub-sample and adjusting it to the real territorial distribution of immigrants*

The HCSO is currently testing different sample frames to choose the most suitable one for getting real territorial subsamples for estimation of number of immigrants. The first and actually ongoing step is to find the right sample frame, using the available administrative sources to design a stratified sample. In this way we could identify small areas or settlements with a high concentration of migrants. A second step could be to supplement the usual LFS sample with sample units deriving from another sample frame suitable for measuring immigration. However, cost of the survey on a boosted sample will be significantly higher than in the case of the basic survey. It is very important to note that improving the available administrative registers as described in Section 3.1. is crucial for the optimization of the sample-frame applied.

• *Steps towards the harmonization of the definition of household membership*

The Hungarian LFS is fully harmonized to the ILO recommendations, Eurostat legal acts and regulations. The HLFS sample includes only those households where one member of household has been working abroad for no longer than one year. The definition is written in the handbook for interviewers, however (as we have learned from the SEEMIG pilot study) the definition of this kind of outworkers should be more precise if we would like to analyze migration process more in-depth. In the future more practical exercises will be introduced in the training program of interviewers to help them better understanding the definitions written in the LFS handbook.

• *Steps towards the standardization of concepts for identifying foreign-born population*

Concepts are already standardized.

• *Steps towards the inclusion of supplementary questions regarding emigration or labour force out-migration*

Better understanding of labour force outmigration is of high priority in the Hungarian LFS. We have already had a special ad hoc module of the LFS on emigration's intentions, and this module gets repeated from time to time. Also, exploratory work is being done to improve current batteries regarding employment-migration and the possibility of adding a regular additional battery to cover the issue of emigration is also under investigation. See also "*Possibilities of using SEEMIG pilot methods (formulating new, constant LFS questions concerning outmigration)*" below.

•*Steps towards the harmonization and wider selection of non-response codes, in order to identify when the whole household went abroad.*

As part of the ongoing project for improving the methodology of LFS, we plan a multi-level non-response code system, which will help us getting a more detailed picture on the reasons of non-responses due to language barriers. Also, the inclusion of additional codes to better identify the reason for the absence of the household is under consideration (e. g. members of household moved abroad). It would be desirable to harmonize a migration-sensitive non-response code system on the EU-level.

- ***Possibilities of using SEEMIG pilot methods (formulating new, constant LFS questions concerning outmigration)***

On the basis of experiences earned in the SEEMIG-LFS pilot study it was concluded that being a nationally representative survey with a large sample-size, the LFS has the potential to serve as a basis for a reliable estimate of the number of emigrants coming from a country. We propose to attach a special battery to the LFS, but only on a time to time basis.

According to our proposal, based on the SEEMIG pilot study, each household in the LFS sample, should be asked whether or not (a) a current household member, (b) a former household member (someone who left more than a year ago), (c) a sibling of any of the household members, (d) a child of any of the household members is currently living abroad. The distributions of the answers given to these questions will then serve as a basis for an estimate of the emigrant stock data in a given year. To calculate this estimation, we advise to use the "Generalised Weight Share Method" to handle the indirect nature of the data collected. Further refinements to the estimates can be made by taking into account the information on the *time of emigration* as well as the *temporal patterns of the stay* (permanent stay or some form of commuting). These data are also collected from the LFS household-members.

Besides estimating the mere number of emigrants, the battery should also aim at collecting some basic demographic and social data on them. Although substantial sample-attrition can be expected at this stage of the study, the methodology can still provide unique information on the composition of the stock of emigrants on a yearly basis. It is very important however to limit the battery only to the most important set of data. We recommend therefore to collect information on the followings:

- gender
- year of birth
- highest level of formal education
- current labour market situation
- country of stay

A detailed description of the methodology proposed will be given in the upcoming SEEMIG report on the second phase of the pilot study.

The LFS is a very overload survey with an extensive standard block and also regular extensions. At the same time it is the most costly survey carried out in the Office, therefore finding additional resources for additional tasks is certainly an issue.

3.3. Migrant-specific surveys

Survey research offers great potentials to generate data in the field of migration not only to supplement but also to substitute administrative and statistical data. Relatively low costs and flexibility make survey techniques especially attractive. Due to certain special characteristics of the migrant population (rare and hidden population) however, exceptional challenges occur when sampling is concerned.

• *Introducing systematic, internationally standardized, survey-based migration-related data collection*

To follow emigrants from the individual countries (as described above) we propose a specific module to be linked to LFS from year to year. This module should serve as a basis for producing a time-series of the emigrants coming from Hungary and also to provide some basic statistical information on their social and demographic characteristics. As LFS is highly standardized across Europe and because the methodology suggested is also easy to standardize, the described module could form part of the core module of LFS. In this way internationally comparable data on the magnitude of emigration as well as on some key characteristics of the migrant population could be produced across the European countries.

To collect more in-depth information directly from members of this population in a representative manner however we propose to launch a new harmonised immigrant survey in the main immigration countries and also in European countries with ethnically homogenous immigrant groups. The main motivation for this is that it is easier to reach out to immigrants in the destination country than in their countries of origin. By applying a standard methodology with harmonised survey-questions in each country, this approach would provide the opportunity to exchange data between countries. Valuable data on emigrants for example from Romania could in this way be provided by collecting data from the respective surveys in Italy, Spain, Hungary etc., while in the case of Hungary, data collections in Germany, Austria and Great-Britain could serve as primary data sources.

Finally, smaller-scales cooperation between social research institutes in the region and even across Europe would also be welcome at several levels. For example designing and applying standardized questionnaire-batteries in the various electronic emigrant-surveys appearing on various internet sites would be of great added value. As research institutes as well as civil organizations and other bodies frequently publish electronic surveys targeting international migrants on their websites, comparability of some of the data collected this way would be desirable. The questionnaire developed in the framework of SEEMIG pilot study 2. could certainly serve as a good starting point for a transnationally applicable module. Targeted surveys focusing on special forms of migration (such as emigrants from one country to another, commuting, cross-border migration etc.) also require transnational cooperation between two countries.

•*Designing surveys at transnational, national and on a local level*

As far as national and local surveys are concerned, in the field of *immigration* studies a relatively rich set of cross-sectional survey data is available in Hungary. Since the political transformation, several focused surveys have been carried out on ethnic Hungarian immigrants from the neighbouring countries and also non-Hungarian emigrants have been extensively covered. The current improvements of LFS (as described before) also target these groups. Statistical data of immigrants are also available from the Census. Moreover, incompletion of administrative registers on immigrants in Hungary constitutes a significant obstacle to create a representative sample of this group in Hungary. Therefore we propose not to carry out any large-scale representative survey on immigrants until the necessary improvements of these registers as described in the section on Administrative Data sources. Even after that, when it would be possible to have a representative sample, we suggest to consider longitudinal rather than cross-sectional surveys in the field of immigration in Hungary, although we are fully aware that this creates great methodological difficulties and it is more costly .

Emigration receives markedly more attention in Hungary not at least because of its increasing magnitude. Reliable, systematic data is badly needed in this field – to improve estimates on the extent of out-migration as well as to gain additional knowledge on the social and demographic composition of the emigrants, and also to understand their motivations and experiences in their destination country. The SEEMIG-LFS pilot study has been a great step forward in this context providing valuable direct as well as indirect experiences in the field especially with regard to the composition of the out-migrant population. It is necessary to carry out further analyses of the pilot survey comparing it systematically to other data sources. The action plan foresees immediate steps even within the life span of the SEEMIG project itself.

To obtain more in-depth information about emigration and especially to get information directly from the migrants, other approaches are needed. Instead of aiming to have national representativeness we propose more focused studies dealing with specific groups of migrants as mass migration includes groups among whom social differences are extremely wide. Specific surveys on various age groups could also be relevant, most importantly targeting younger age groups, which have specific problems in accessing labour markets, forming families. The Hungarian SEEMIG team will consider project opportunities to carry out such surveys. Studies on emigrants from smaller localities (e.g. typical sending towns) can for example apply the SEEMIG pilot method to build a representative dataset of emigrants from the locality and then to approach the emigrants directly. An alternative to the SEEMIG way to ensure representativeness of the emigrant sample is applying Respondent Driven Sampling (RDS) – a method already successfully used in a study of Hungarian migrants to Austria. Choosing RDS would be also useful in other important areas, such as emigration of members of certain professional groups (e.g. health-care professionals, construction workers etc.) and also to survey issues such as families split by migration etc. In research fields, where representativeness is less crucial well-designed qualitative studies could also help to fulfil information-needs.

Besides providing more insight into the actual process of immigration and emigration, survey research can also be useful in helping to understand the wider social context of migration processes. While surveys on migration intentions as well as surveys on attitudes towards immigrants are carried out in Hungary on a more or less regular basis, further research would be needed on the social values and attitudes concerning current emigration processes in Hungary as well as on the emigrant communities abroad. Research in this field could also help us to better understand (and possibly to handle) the notable social distrust experienced during the SEEMIG pilot study around the issue of international migration.

•*Possibilities of planning modules for already existing international surveys (for instance related to outmigration)*

Considering the genuinely international nature of migration so far it has received relatively little attention in the standardized international value-surveys such as ESS and ISSP. It was more than a decade ago that relevant, transnationally comparable attitude surveys were last carried out in the field. (a survey on immigration in ESS in 2002, and PPA surveys in 2000-2003. Including a module

focusing on social attitudes towards the current intensity of worldwide migration in one or the other of these standardized surveys is highly recommended.

• Steps towards setting up longitudinal migration relevant survey

As described above, a longitudinal survey of immigrants in Hungary (either on the national level or in areas with a high level of immigrant-concentration) could extend our understanding of the immigration process to a notable extent. By following members of these groups over a longer period of time, the process of integration could also be studied serving not only academic interest only but also supporting policymaking.

• Possibilities of conducting national longitudinal surveys based on local government questionnaires

An innovative method has been introduced to measure certain forms of migration which cannot be covered by national statistics: the TÁRKI Local Government Monitoring and Database Project (LGMDP). Migration related topics covered by the waves of LGMDP were informal economy, foreign labour, suitcase trade; refugee integration; migrants on the local labour market and emigration. LGMDP surveys while providing additional municipality level information (e.g. ethnic composition, commuting migrant workers, open-air marketplaces) that other existing databases do not contain, its data can be pooled with all other municipality level statistics. This large set of variables gives an abundance of explanatory power to the models explaining the causes and consequences of the migration process. The data cover parts of the migrant population which cannot be reached by standard statistical methods (illegal migrants, commuting workers, shuttle traders) as well as some aspects of the migration process which are too specific to be properly analyzed by macro-statistical methods (such as integration, FDI on local level, wages on the local labour market, spread of open-air markets, etc.). Methodologically the survey is based on questionnaires sent to all Hungarian municipalities. Data from the longitudinal panel's waves are pooled into a cumulative database which enables users to make time series analysis as well projections for the future. The research design allows for a quick and up-to-date analysis (the time required for the fieldwork is about two or three months) and is inexpensive since it is based on the semi-voluntary participation of the respondents at local governments.

• Further remarks

As the intensity of emigration from Hungary is increasing and emigration is mainly motivated by better employment prospects and higher level of income in the destination countries, we expect that remittances sent home by the migrants are gaining increasing importance in the Hungarian economy. Hard evidence on these is however very scarce. To better understand not only the demographic and labour-market nature of migration but also its wider economic consequences, targeted household surveys including detailed questions on remittances received are needed. Also, inclusion of

remittance-items in the already existing household budget survey and also in EU-SILC should be considered. The review of the possibilities will be done by a special group within HCSO with the inclusion of outside experts.

As massive emigration is a relatively new phenomena in Hungary, migration-related questions should be included in the already existing larger social surveys to an increasing extent. Depending on the nature of the specific survey, in each case it should be considered whether it would be appropriate for accommodating certain items with regard to migration. For example would a consecutive wave of the GGP survey (Gender and Generation Programme's longitudinal survey) in Hungary be carried out, we would propose that migration (emigration) becomes a priority among the various demographic domains covered.

3.4. Census

All proposals below will be put forward in the professional team building up the new census for 2021 within HCSO.

•Steps towards the harmonization of questions referring to migration history

Questions referring to migration history were included in the census in compliance with *EU Regulation (EC) No 763/2008*:

- place (country) of residence at birth

- former usual residence and date of arrival to the present place (country)

- former place of residence abroad (if having resided abroad for at least one year), date of arrival (or return) in the country

Based on the above, no harmonization is needed in Hungary in this area.

•Steps towards the inclusion supplementary questions regarding immigration

Besides compulsory elements prescribed by *EU Regulation (EC) No 763/2008*, member states are free to include additional questions in the census questionnaire. In case of Hungary – due to the deficiencies of migratory data coming from administrative data sources and therefore the limited possibilities of analysis – most of the migration-related questions recommended by the SEEMIG Synthesis report (*Analysis of existing migratory data production systems and major data sources in eight South-East European countries)* are indeed necessary and strongly recommended to be included in the next census. The recommended questions are as follows:

- *legal status* of third-country nationals at the time of their first arrival with aims for longer term residence and at the time of the census (this is important to be able to identify persons with various migrant background – e.g.: refugee background – in order to evaluate various factors of the integrations process)

- *date of acquisition of Hungarian citizenship* and *the type of naturalization* – in case of those who have not had a Hungarian citizenship since their birth (this is especially important in case of Hungary, since the number of new citizens is relatively high)

- *place (country) of birth or former citizenship of parents* – in case of the whole resident population (this is important to be able to identify second generation migrants – those who were born in Hungary and/or are Hungarian citizens, but whose parents are immigrants).

In case of persons with dual or multiple citizenship, the Hungarian census registered both citizenships, in cases when one of the citizenships of the person was Hungarian and also when both citizenships were other than Hungarian.

Out of the new questions recommended above, *place of birth of parents*, *date of acquiring citizenship* and the *type of naturalization* was included in the recommendations of the UN for the 2010 census rounds as 'not core topics'.

- ***Enhancing the successful enumeration of foreigners living in Hungary***

In order to reach a more successful enumeration of foreigners living in Hungary, it is necessary to further improve foreign language questionnaires and foreign-language enumeration procedures. Based on experiences of the 2011 census and foreign-language LFS, access rates to foreigners need to be improved. Language training of interviewers or employing new interviewers with already sufficient language knowledge does not seem to be a realistic goal, given the already 40 000 interviewer staff. On the other hand, questionnaires and information material in foreign languages could prove to be effective. There is a need to increase the number of foreign languages to which the census is translated. Online data collection is envisaged.

In the 2011 census – similarly to the previous one ten years before – foreigners were provided with foreign-language questionnaires. The online questionnaire introduced in 2011 was also available in English, while paper-based questionnaires were available in the languages spoken by major immigrant groups (English, French, Chinese, Vietnamese) as well as in the language of nationalities or recognized ethnic groups living in Hungary (Armenian, Bea-Roma, Bulgarian, Croatian, German, Greek, Lovari-Roma, Polish, Romanian, Russian, Rusyn, Serbian, Slovak, Slovenian, Ukrainian). The evaluation of this method is hindered by the fact that all questionnaires completed online were received through one channel and they were not sorted out by language. Moreover, only Hungarian paper-based questionnaires could be processed in an automatic way. The content of questionnaires completed in other languages had to be copied before being processed. Therefore, no information is available about the number of paper-based questionnaires completed in languages other than Hungarian. One package of foreign-language questionnaires per enumeration district (that is, one questionnaire in each above mentioned language) proved to be sufficient in most cases, meaning that data provided by respondents not speaking Hungarian were recorded on a Hungarian questionnaire, but the interview was conducted based on a questionnaire in a foreign language. Interviewers also reported that some foreign respondents provided answers with the help of a household member (e.g.: child) speaking Hungarian. There is a need to organize special trainings for the enumerators.

- ***Steps towards enhancing data on emigration and inclusion of supplementary questions regarding out-migration and return migration.***

Regarding out-migration, partial information can be gathered from the 2011 census, in particular from the questions asking the following:

- the number (from the dwelling questionnaire) and characteristics (from the personal questionnaire) of persons staying abroad on a temporary basis (less than a year);

- the number of persons staying abroad on a permanent basis (living or expected to live abroad for at least a year). This information was gathered from the dwelling questionnaires, no personal questionnaire had to be completed about these persons.

- the previous place of residence abroad (if the person stayed abroad for at least a year continuously) and the date of return;

- the place (settlement) of employment: foreign country could be selected.

The first two questions attempted to register those staying abroad at the time of the census, the third question tried to track persons returning after a minimum stay of one year abroad, while the fourth question referred to persons having their usual residence in Hungary, but working abroad (commuters).

No information was gathered in the census about the characteristics of persons permanently staying abroad. Data providers had been asked to give the number of persons staying permanently abroad out of those having residence in the dwelling. In case the data provider had no such information or the dwelling was unoccupied the persons staying permanently abroad were not covered. Similarly, only a partial group of returnees were grasped by the census: those returning home after a minimum stay of one year abroad.

Taking into account all the above mentioned deficiencies, as well as the trend of increasing out-migration in Hungary, the following steps are suggested for consideration:

- the introduction of a separate questionnaire (or the inclusion of a question block) in the census for those having a permanent place of residence in Hungary and living or working abroad (irrespective of the length of stay or employment) in order to register – besides basic demographic characteristics – the *country of destination*, the *starting date of stay/employment abroad*, the *expected length of stay/employment abroad* (less or more than a year), the *reasons for staying/working abroad*, *type of employment abroad*. The questionnaire will be partially based on the questionnaire developed for the SEEMIG pilots.

- the inclusion of a separate question block referring to those not having a permanent place of residence in Hungary but who have emigrated from the responding household (in the last X of years);

- the inclusion of further complementary questions in case of former place of residence abroad: *how many times has the person lived abroad*, *which country/countries has the person lived in*, altogether *how long the person lived abroad*, in case of many occasions: the *length of last stay abroad*, the *motivation of staying abroad*, *date of last return*, the *reason for return*;

- the development of a methodological best practice to identify complete households moving abroad.

- ***Preparations to introduce register-based census***

Despite its many advantages (lower costs, members of the population are free of the burden of data provision, faster data procession, the possibility of more frequent enumeration, e.g.: every five years), the introduction of register-based census is far from being an easy step to make, since many of its preconditions require a long preparatory phase.

Steps to be taken:

- clarify the possibilities and conditions for the introduction of register-based census in Hungary (e.g.: ways to handle the problems arising from the difference between registered and actual places of residence);

- define migration-related census data that can be derived from existing registers at the moment and explore to what extent these data cover the variables prescribed by the EU;

- reveal the possible directions for the development of registers relevant for migratory data (with respect to the provision and control of necessary basic data);

- define legal preconditions for the introduction of register-based census, with respect to interlinking various administrative data sources (applying common personal identification numbers) and the improvement of these data content.[3]

From the point of view of migration, however, – considering the deficiencies of migration-related administrative data sources – register-based census do not necessarily mean a step forward, since traditional census is already richer in migration-related data (containing data not included in registers). It is important to assess whether migration-related registers can be developed/improved in a way that besides presently prescribed (compulsory) migration-related data, additional data recommended in the previous point could also be accessed.

- ***Comparable EU census data about migration - Census Hub***

EUROSTAT has developed a new method to collect and disseminate 2011 census data of the member states. Census data previously transmitted in simple .xls files are now disseminated in multi-dimension data hypercubes. The sixty hypercubes, most of which consists of eight dimensions, are stored by the member states and the CENSUS HUB system makes it possible for the public to make prompt data queries from the EUROSTAT website. By the end of March 2014 deadline set by the EU Census Regulation, census data of the member states, mostly on regional level, has been made available on the EUROSTAT website.

In the fifty hypercubes containing population data, from census topics that are most relevant in relation to migration, country of citizenship is available in 34; country of birth in 26; place of residence one year prior the census in 10; year of arrival in the country in 8; the fact of arrival in the country since 2000 in four hypercubes, allowing to observe migration within the EU.

---

[3] It is also recommended to consider the key messages of presentations given at the workshop entitled *'Statistical use of administrative datasets – with special respect to the census'* held on 29 March 2010 at the *Fényes Elek Műhely,* a workshop of the Hungarian Central Statistical Office. (http://fenyeselekegyesulet.wordpress.com/eloadasok/2010-2/adminisztrativ-adatallomanyok-statisztikai-celbol-valo-hasznositasa/, in Hungarian)

The fact that the census hub system is a development for EUROSTAT without antecedent gives the launch a certain risk, but if the years of preparation end successfully then this new, network- based pull method way of dynamic dissemination is planned to be used in other statistical fields as well.

3.5. Other

- *Remittances*

One of the consequences of the international migration of the labour force are the remittances sent by migrant workers. On one hand remittances support the family members left at home, on the other hand they are an important part of the national economy. The analysis of in-and outflows of money transfers at household level can serve as a substantial part in understanding and refining migration flows. So far these data were hardly utilized in Hungarian migration statistics. The methodology of the integration of remittances data into migration statistics must be elaborated as well as the standard procedures of the transfer of such data to the HCSO. The methodology of estimating remittances should also be revised and better integrated into migration statistics. There are major methodological problems concerning the estimation of migration stocks living abroad.

- *Big data*

Call logs, mobile bank transfers, text produced by online users (blog entries, social media entries), online queries, satellite shots, images, videos, etc. all fall into the category of big data. New information can be gathered from these by phishing, i.e. looking for interesting and valuable connections among items of this mass of data. Special IT equipment is necessary to convert these data into usable pieces of information. With respect to migration, the volume of daily (cross-borders) commuting, attitudes towards migrants and migration could be observed with the help of 'big data' data sources.

'Big data' data sources are usually characterised by their immense quantity, frequent production and changing composition. They are most often unstructured data which need innovative processing methods in order to be able to have an overview and new utilisation of them.

The unstructured nature of big data means that connection between data is previously unknown. They mainly provide opportunity to learn more about people's habits and behavioural patterns in more detail. This, on the one hand, generates commercial opportunities in the private sector, but can be considered a valuable input for official statistics both on its own or combined with traditional data sources.

Big data could be invaluable for statistics due to the fact that on one hand they show actual activities of the persons examined (contrary to what they report about themselves), and they provide information in real/actual time on the other, which make these data more complete and valuable over time.

Big data is increasingly attractive for the statistical utilization and the HCSO has also started to examine their potential. After HCSO will define the general conditions of using this type of data sources, it will be worth to develop a clear methodology to use them to investigate migration-related issues.

4. LOCAL LEVEL ACTIVITIES FOR MIGRATION DATA IMPROVEMENTS

4.1. Setting up local developmental databases

The existence of a local monitoring system and database on migration, demography and human capital is essential for evidence based local strategy making. The integrated accessibility of data, their continuous updating, and in case of certain types of data, their presence on spatial information system are vital elements of planning, monitoring and evaluation. The local databank channels information from various data systems: census, register based data, survey data, predictions, population estimations, foresight scenarios.

4.1.1. Census

Census data have a special value for cities and villages in general and for the city of Pécs in particular. In spite of the fact that data is only updated every ten years, this is the SOLE source of data that lends itself to providing data on the settlement as a whole as well as on practically any territorial level. Census data integrates data from areas such as educational, labour market, demographic, internal migration data of individuals, composition of households and characteristics of dwellings. A major threat is posed by the fact that the alternative of a full coverage census, a fully or partially register based census is being discussed, a move that would deprive settlements of a fundamental database and would essentially prevent evidence based strategic planning in an area based manner. A fully or partially register based census namely, would transform settlements into one of the variables of datasets on individuals: the full coverage of settlements would be abolished, especially as regards households and dwellings.

4.1.2. Register based data

The local databank can collect register based data from various sources: there is a great variety as to data producers, the nature of data and the persons about whom data is collected. The HCSO registers basic population data (births, deaths, migration), COAEPS deals with residence, the OIN registers foreigners, the NHIFA the data of the insured, the NTCA collects taxes and the data of taxpayers, National Employment Service registers the unemployed and work permits for foreigners. The database of the University of Pécs, ETR registers data on students studying at the university, foreign and Hungarian. As a recommendation, it would be advisable that members states of the European Union register in their work permit or other registers not only citizenship and country of origin but also former residence, including settlement: this would make it possible for European cities and villages to determine and manage their migration networks.

The part of the local databank that would incorporate register based data, can rely on the planned migration databank of the HCSO. This integrated statistical (not administrative) database organises data from various register based data sources by way of connecting individual records via identification keys, without personal data. As part of the local data strategy, cooperation is suggested with experts of the HCSO to introduce the dimension of settlements (and thereby counties) into the design of the dataset. At the same time low investment access is also to be provided to the databank,

chiefly for operators of local databanks so that regular updates can be made to local databanks, by default in the research labs of the HCSO.

The HCSO should obtain data from the National Health Insurance Fund on persons who did not pay social insurance for a longer period (e.g. one year). Not only a national but a local statistical survey could also be carried out on all of these persons in the settlement, checking whether they are emigrants or have other reasons for not paying social insurance.

4.1.3. Survey data

The local data strategy covers the idea of archiving prior local survey data as well as the objective of starting a new epoch: in the future, surveys on the urban society, its labour market shall be available in the local database in terms of their design, results, including (preferably) the datasheets as well. National surveys using large samples do not lend themselves to direct interpretation of local social processes either in case of LFS or EU-SILC or other large national sample surveys. This is especially true of question panels of large sample surveys that produce a small number of cases, as in the case of migration related questions (indicated by the SEEMIG WP4 Hungarian country report *Analysis of existing migratory data production systems and major data sources in Hungary* as well). At the same time when time survey results are put into the context of regions and big cities as a whole, they provide a useful set of references for the interpretation of other types of data included into the local databank. For this sociology department puts together ideas?

4.1.4. Estimations

Besides conventional statistical methods, hard-to-grasp phenomena like migration made it necessary to develop the method of migration estimation based on informal data gathered among local informed actors. TÁRKI (an independent research institute) has been conducting a longitudinal research among Hungarian local authorities on their economic and social circumstances by surveying local actors. Several waves of the research (1995, 1996, 1997, 1998, 2000, 2002, 2005) dealt with migration beside other themes such as local informal economy and labour market, including the expectations and opinions of local actors about the trends under scrutiny. It is advisable that the Municipality of Pécs joins the research since the resulting data, interpreted in the context of other data from the local database can be of great service for local strategic planning, whereas data from the survey of a particular year can be interpreted in the context of other Hungarian cities. Data from these local estimates are to be uploaded into the local databank.

4.1.5. Population projections, foresights

Based on prior trends, using standard methods, statistical population predictions are produced among the activities of WP5 of SEEMIG. Projections for the national level will be accompanied by projections for all counties as well as for the city of Pécs. Results are to be uploaded into the local databank.

Beside statistical population predictions, there are qualitative methods for the prediction of future trends producing scenarios. Among these foresight scenarios stand out, a method that is applied in WP5 of SEEMIG. It is advisable to apply this on the local level also so that not only conventional data but also local expert as well as local lay knowledge is channelled into local strategic planning.

4.2. Capacity building at local governments

The setting up of the databank and the development of its virtual home requires IT development. The host of such a development and the databank can be the Municipality of Pécs, the Pécs Urban Development Agency or VÁTI in the framework of a pilot project. A spatial informatics alternative is offered by the option of developing the databank as a series of layers to the already existing urban spatial informatics system created and maintained by Pécsi Geodéziai Ltd.

It would be advisable that future maintainers of the databank have an insider view on its design, uploading of census and register based data, the entry of survey data, locating of other types of datasets. This element of capacity building can be tailored to concrete person(s) in a learning-by-doing manner during the time of the development of the databank.

Databank users have to be supported in the form of special training programs which take note of typical user needs and competencies. Without generating and updating the skills of experts around decision makers, it would be futile to develop the databank in the first place, since the numbers and other figures contained in it are 'dead letters' only if they are not interpreted continuously, applying the conclusions supported by these interpretations into practical life. Similarly, those professionals who regularly take part in the creation of urban strategies but are not employed by the Municipality are advised to develop the skills necessary for using the databank. National research institutes and HCSO included in this.

For actors of higher education or research (lecturers, researchers, students), skills for using the databank could be acquired via a peer tutoring system, making databank skills one of the prime element of local actors' mutual learning processes.

5. IMPLEMENTATION POSSIBILITIES

Activity		2015				2016				2017				2018				2019			
		Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
Database	SEEMIG	██	██	██	██	██	██	██	██	██	██	██	██	██	██	██	██	██	██	██	██
	National					██	██	██	██	██	██	██	██	██	██	██	██	██	██	██	██
	Local																				
National	Admin									██	██	██	██	██	██	██	██	██	██	██	██
	LFS			██	██																
	Survey					██	██	██	██	██	██	██	██	██	██	██	██	██	██	██	██
	Census																	██	██	██	
	Other																				
Local						██	██	██	██	██	██	██	██	██	██	██	██	██	██	██	██

Meetings with experts and/or stakeholders from the relevant institutions

Please take note of the expert interviews and/or meetings that you had during the drafting process of the Action Plan (between November 2013 and January 2014)

- **9 December 2013**
  HCSO, DRI, Dr Endre Sik (Eötvös Loránd University, TÁRKI)

- **20 February 2014**
  HCSO, DRI, Dr Judit Tóth (University of Szeged)

6. PLAN FOR SUSTAINABILITY

HCSO as SEEMIG's Lead Partner is committed to provide the sustainability of SEEMIG's proposals. The action plan will be integrated into the development strategy of HCSO. Negotiations will also be made to integrate the action plan into the national migration strategy, to gain the status of an approved policy document as the enforced cooperation of various partners is required. Further institutional steps will be defined soon.