

## SURVEY OF HISTORICAL DATABASES

### WITH LONGITUDINAL MICRO-DATA

#### The second questionnaire

For more information about this questionnaire or questions about entering specific information, please contact Kees Mandemakers ([kma@iisg.nl](mailto:kma@iisg.nl)) and/or Tatiana Moisseenko ([tatiana.moisseenko@iisg.nl](mailto:tatiana.moisseenko@iisg.nl))

#### Table of Contents:

- |       |   |   |
|-------|---|---|
| I.    | A | General (identifying) information about databases |
| II.   | A | Contact information                               |
| III.  | A | Sources: core characteristics                     |
| IV.   | A | Database: core characteristics                    |
| V.    | A | Publications and reports                          |
| VI.   | B | Observations                                      |
| VII.  | B | Sampling design and procedures                    |
| VIII. | B | Data collection                                   |
| IX.   | B | Linkage process                                   |
| X.    | C | Sources: the main characteristics of every source |

The questionnaire comprises three sections:

**Section A** includes the questions related to the most general and important information identifying the content, scope and provenance of the databases and the information about their creators.

**Section B** contains more specific and detailed questions about databases, such as the period(s) of observation, sampling design and procedures, data collection, linkage process and others.

**Section C** contains detailed questions about sources used for the databases: their type, scope, content, state of preservation, etc.

#### Section A

##### I. General (identifying) information about the database

<b>1. Title of the database</b>	Historical Sample Portuguese Social Mobility
<b>1.a. Subtitle</b> , which brings meaning to the title (scope, place, time period):	1850-1960
<b>2. Abbreviation</b>	HSPM

<b>3. Links to website(s):</b>	
3.a. Homepage	<a href="http://www.mobilidadesocial.uevora.pt/">http://www.mobilidadesocial.uevora.pt/</a>
3.b. Get to data	

<b>4. Abstract:</b> describes content of the database. Max. length: 300 words Please indicate: <ul style="list-style-type: none"><li>◦ Scope and main goal</li><li>◦ Time and territory covered by data</li><li>◦ Sample strategy</li><li>◦ Main sources</li></ul>	<i>Original goal</i> To study national and regional patterns of social mobility, using marriages registers and Hisco coding scheme for international comparison. <i>Sample definition</i> The sample was defined by common time cohorts for different national contexts chosen from different social and economic patterns revealed by previous monographic academic studies. Data about the bridal pair, occupations and geographic location of all participants (including testimonies) in each marriage record (45,723 records).
---	---

	<p><i>Geographic area under observation:</i> Alentejo region: Évora (Urban/rural parishes) and Portalegre (Urban/rural/industrial); Setúbal (Urban/rural parishes), coastal town south of Portugal; Barreiro (industrial ‘village’) near Lisbon; Figueira da Foz (Urban/rural parishes), in the Centre (coast); Lisbon (Urban); Oporto (Urban); North Interior: Viseu (rural/urban); Centre Interior: Covilhã (rural/ urban – industrial textile); Atlantic Islands: Azores - Ribeira Grande (rural, proto-industrial), Ponta Delgada (urban/rural)); Madeira: Funchal (urban/rural).</p> <p><i>Main sources:</i> Marriage records (parish registers before 1910; civil registers after 1910)</p>
--	---

<p><b>5. Keywords:</b> Please use the recommended keywords if they are applicable: <i>demography, life course, census, church register, civil certificates, population register, history, social science, genetics, migration, occupations.</i></p>	Civil certificates, church register, history, social science, migration, occupations
Please add your own keywords, if you have data not covered by the recommended terms.	Portugal (19 <sup>th</sup> to 20 <sup>th</sup> centuries)

<p><b>6. Citation:</b> Indicate how you want others to cite your database.</p>	Fonseca, H.F.; Guimarães, P.E. - <i>HSP-SM. Historical Sample of Portugal on Social Mobility (1860-1960) – Version 2.0. – Évora: University of Évora, 2014</i>
--	--

<p><b>7. IDS compatible:</b> Indicate with <i>Yes</i> or <i>No</i> whether the database is IDS compatible, if <i>Yes</i>, please specify.</p>	No (individual names are not available due to privacy data protection)
---	---

<p><b>8. Has the database already been completed or it is still under construction?</b></p>	Under construction. Data collection is completed. Present status: coding geographic names in all marriage registers (realized 60%).
8.a. If completed, please indicate the years of its construction?	7 years
8.b. If under construction, please indicate, when it is planned to complete it?	2016 (December)
8.c. Please add a brief description of future plans for the database.	After publishing its detailed description of the database and the methodology used, we will be presenting analysis of mobility tables and other data, and make the database available to the academic community.

## II. Contact information

<b>1. Name of institute or organisation</b>	CICP - CENTRO DE INVESTIGAÇÃO EM CIÊNCIA POLÍTICA / University of Évora
1.a. Website	
1.b. Location: city, country	Évora, Portugal
1.c. Postal address	CICP - Centro de Investigação em Ciência Política Universidade de Évora Palácio do Vimioso Largo Marquês de Marialva, 8 7000-809 Évora Portugal
1.d. Phone	+351.266706581

<b>2. Name of primary responsible person</b>	Dr. Paulo Eduardo Guimarães
2.a. His/her email address	<a href="mailto:peg@uevora.pt">peg@uevora.pt</a>
2.b. Postal address	Departamento de História Universidade de Évora Colégio do Espírito Santo Largo dos Colegiais, 2 7000-803 Évora Portugal
2.c. Phone	+351.309902574

<b>3. Administrative information</b>	
3.a. When this form was filled?	Setúbal, 17 <sup>th</sup> February 2015
3.b. Who did it?	Paulo Guimarães

<b>4. Main economic funding</b> (Name of organization(s) who made the grants /sustain it)	Portuguese Science Foundation = Fundação para a Ciência e a Tecnologia. Project PTDC/HISHIS/110827/2009
---	---

### III. Sources: core characteristics

#### 1. Type of the sources.

Indicate how many sources were used for the database and what kind (register, census, certificates ...). Please enter *Yes* or *No* and the time period for the main sources. In case of other sources, not listed below, please add their type and specify their main characteristics. Detailed questions about the characteristics of all core sources are in section C.

	<i>Type of source</i>	<i>Yes/No</i>	<i>Start year</i>	<i>End year</i>	<i>Explanations:</i>
1.	Baptisms	N			
2.	Marriages from church registers	Y	1860	1910	After the Republican Revolution of October 1910, the civil registration was introduced in Portugal
3.	Burials	N			
4.	Population registers, maintained by church or state	N			
5.	Civil birth certificates	N			
6.	Civil marriage certificates	Y	1911	1957	After this date there is no information in the source about the parents' occupation of the bridegrooms.
7.	Civil death certificates	N			
8.	Population Census	N			
9.	Nominative lists	N			
10.	Military draft records	N			
11.	Other:				

### IV. The database: core characteristics

<b>1. Period covered by data:</b> give first and last year of date, if possible	1860 - 1960
<b>2. Territory covered by data</b>	Alentejo region: Évora (Urban/rural parishes) and Portalegre (Urban/rural/industrial); Setúbal (Urban/rural parishes), coastal town south of Portugal; Barreiro (industrial 'village') near Lisbon; Figueira da Foz (Urban/rural parishes), in the Centre (coast); Lisbon (Urban); Oporto (Urban); North Interior: Viseu (rural/urban); Centre Interior: Covilhã

	(rural/ urban – industrial textile); Atlantic Islands: Azores - Ribeira Grande (rural, proto-industrial), Ponta Delgada (urban/rural)); Madeira: Funchal (urban/rural).			
<b>3. Geographical characteristic:</b> local, regional, national, cross-national	cross-national			
<b>4. Units of observation.</b> Please enter <i>Yes</i> or <i>No</i> for each unit, which forms the sample, the number of units and write explanations/comments. Add other units if they are not listed below, for them explanations are especially important.				
	<b>Units of observation:</b>	<b>Yes /No</b>	<b>Number of units</b>	<b>Explanations:</b>
1.	Individuals	Y	365,784	Individual names were not recorded in the database. Some entries do not have information about occupation.
2.	Married couples	Y	45,723	Idem
3.	Families	N		Idem
4.	Households	N		Idem
5.	Farms	N		Idem
6.	Institutions	N		Idem
7.	Other: Occupations	Y	3,430	Portuguese occupation titles.

<b>5. Variables per unit included in the database</b>	
<u>On individuals:</u> <i>Data of birth and dead, age, gender, marital status, religion, occupation, migration, relationship, etc.</i> Please add more variables, if they are not in the list	The sample includes the occupations of the bridal pair, its geographic origin, residence (address). The same data was gathered from the parents and testimonies.  The variables recorded in each marriage book vary according to the role of the individual in the marriage and to the legislation. For all participants we have these variables: source, location, date of the marriage, civil status, age, occupation, residence, place of birth.
<u>On households:</u> <i>Type of household, children present, age and number of children, etc.</i> Please add more variables, if they are not in the list	

<b>6. Kinship relations:</b>	
6.a. How is kinship recorded in the database?	By specific fields in the database
6.b. How deep (number of generations) is kinship information going?	One (father/mother to -son/daughter)

<b>7. Completeness</b>	
7.a. Are all variables from the sources included in the database?	No, names, notes, signatures, annotations and other random information in records were not registered
7.b. Are all individuals who lived in the households of the sample recorded?	Not applied
7.c. Are the main variables comparable?	The main variable (occupations) is comparable but missing data is a problem in the sources prior to 1911

<b>8. Current data representation:</b> Database Software (e.g. MySQL, MsSql, Access, please specify)	Microsoft Access (2007)
---	-------------------------

<b>9. Access conditions:</b>	
9.a. How does a user get access to the database?	After 2017 by request send to Paulo Eduardo Guimarães (e-mail: <a href="mailto:peg@uevora.pt">peg@uevora.pt</a> )

9.b. What are the conditions and restrictions?	To be defined by the Portuguese Foundation for Science and Technology (FCT).

## V. Publications and reports

### 1. Main publications about the database itself (max. 5)

- Paulo Eduardo Guimarães - *Explorações sobre a mobilidade social em Portugal(1860-1960). Anexo I. Caracterização da AHP-MS V 2.0*. Progress report nº 2 of the project: PTDC/HIS-HIS/110827/2009 (01-03-2011 - 31-08-2013) to the Portuguese Foundation for Science and Technology – FCT, from 01-03-2012 to 31-03-2013. – Évora: NICPRI, 2013. [Detailed description of the database content and methodology, in Portuguese].

### 2. Main or exemplary publications on research based on the database (max. 5)

- Helder Adegar FONSECA and Paulo Eduardo GUIMARÃES (2009), “Portugal e as Sociedades Europeias: Homogamia, fronteiras de classe e revolução sexual (1860-1960)”, H.A. Fonseca e Luis Lobo Fernandes (eds) , A Europa, Portugal e os Desafios da Globalização: Perspectivas Transdisciplinares (nº temático de PERSPECTIVAS, Revista do NICPRI, 2011).
- Helder Adegar FONSECA e Paulo Eduardo GUIMARÃES (2009), “The Social Mobility in Portugal (1860-1960): Operative Issues and Trends”, Continuity and Change, Oxford, 2010.
- Helder Adegar FONSECA and Paulo Eduardo GUIMARÃES (2009), “A Mobilidade Social Intergeracional em Portugal, 1957”, Serrão, J. V.; Pinheiro, M.; Ferreira, M.F. (eds.), Desenvolvimento Económico e Mudança Social, Lisboa, ICS, 2009.

### **Section B**

contains more specific and detailed questions about databases, such as the period(s) of observation, sampling design and procedures, data collection, linkage process and others.

#### **VI. Observations**

1. How do individuals enter observation?	When and only if they married.
2. How do individuals leave observation?	After marriage
3. How do households enter observation?	
4. How do households leave observation?	
5. Are some entry or exit dates unknown?	No
6. Are some entry or exit dates estimated?	No
7. Can observations be linked to geographic locations?	Yes
8. Are the dates and locations of movements within the observation area recorded?	No
9. Are all individuals who lived in selected households recorded? (Selection on basis of the sample or because sampled individuals are living in households)	No
10. Are there related observations that are not included in the database?	The names of all participants and the endorsements on marriage records were not recorded such as the date of death of the wife or husband, for instance, or the date of divorce (after 1911).

#### **VII. Sampling design and procedures: how was sample(s) defined?**

<b>1. Source(s):</b> Which source forms the basis for the sample?	Marriage records (parish registers before 1910; civil registers after 1910)
<b>2. Sampling units:</b> Households, individuals, regions...	Time cohorts and parishes / freguesias, municipalities, regions
<b>3. Variables used for selection:</b> Age, gender, marital status, other	The sample includes the occupations of the bridal pair, its geographic origin, residence (address). The same data was gathered from the parents and testimonies.
<b>4. Selection method:</b> Random, stratified random, total count, clustered, other	The sample was defined by common time cohorts for different national contexts chosen from different social and economic patterns revealed by previous monographic academic studies.

#### **VIII. Data collection**

<b>1. Data collection period:</b> When the data was collected and transcribed?	2009-2014
<b>2. Data collection method:</b> Public digital register, transcription, other	Transcription (typing)
2.a. If transcription, how was the transcription done: <ul style="list-style-type: none"> <li>◦ By individuals</li> <li>◦ From scanned sources</li> <li>◦ From LDS's microfilms</li> <li>◦ Automatic controls</li> </ul>	Directly from the book registers in district archives and in the archives of Civil Registries Offices; and from scanned sources available at the Portuguese archives portal.
2.b. How was the checking of the transcription done? For example, by proof reading?	By proof reading
2.c. When was it done?	After collecting raw data from the sources
2.d. Purpose of the transcription: please indicate <ul style="list-style-type: none"> <li>◦ LDS</li> <li>◦ Research</li> <li>◦ Genealogy</li> </ul>	Mobility research

<b>3. Control methods by researcher:</b> e.g. Internal consistencies such as a death cannot happen before a birth of the same person	The date of the record must be after the date of the marriage; the occupation of the father must be the same that was registered to the godfather if he was the same person
<b>4. Data collection staff:</b> Please indicate the number of people and their position (member of the project, free-lancer, other)	6 researchers (scholars) and 8 research fellows

#### IX. Linkage process

<b>1. Linkage:</b> Which sources and units of observation have been linked: (e.g. birth/baptisms and death/burials...)?	None	
<b>2. Documentation of linking:</b>		
2.a. Programme, manually, ...		
2.b. Name of software if used (and its parameters)		
<b>3. What are the rules for linking?</b> Flags definition (list them: age, name, extra knowledge ...)		
<b>4. How each reconstructed person is traceable to the original sources /transcribed data?</b>	Each database record is coded in order to identify the record source in archives	
<b>5. How is linkage represented in the database?</b> For example, do all occurrences of an individual include a universal identification number (ID)? Or are records linked in another way?	There is a single ID for each marriage record	
<b>6. Linkage percentage</b>		
<b>7. Quality of linkage</b> (own evaluation)		
<b>8. What reference/coding systems have been linked to the data?</b> For example, occupational titles (like HISCO), locations (including geo-referenced systems). Please indicate the name of the system and how it was used. (Yes, No, Partly).	HISCO / Hisclass for occupations and Eurostat geographic coding (Local Administrative Units, LAU)	
<b>Y/N/P</b>	<b>Reference system</b>	<b>Explanations:</b>
Y	Occupational titles:	Occupational titles (HISCO) and Hisclass Occupations were linked to Hisco coding table that was built for Portuguese occupational titles
Y	Locations (including geo-referenced systems):	The Portuguese code for parishes (freguesias)
Y	Religion, civil status etc.:	civil status, legitimacy
	Other:	

### Section C

contains detailed questions about sources used for the databases: their type, scope, content, state of preservation, etc.

Please answer the questions about all the sources used for the database, but do it in a separate form for every type of the source.

#### **X. The main characteristics of the source (per every type of the source)**

<b>1. Official name of the source and its English translation</b>	Marriage registers: church registers recorded in books ('registos paroquiais de casamento') and civil marriage registers ('assentos de casamento'). The marriage records books was created according to the Constitutional Chart (1826) and the Civil Code (after 1867) to register marriages which was considered a civil contract between two persons of different gender for the purpose to create a family. The State guarantees the civil and family rights and duties of individuals, controlling the validity of such unions in the terms of the civil code through these records.
<b>2. Purpose of the source:</b>	
2.a. Why was this source created?	To study stratification and social mobility during the earlier stages of social modernization
2.b. Who created it?	The Portuguese State, through priests (in the case of catholic marriages) until 1910, and through public officials (Registo civil) after the Republican revolution.
<b>3. Scope:</b> What group of the population was documented in this source?	The Portuguese (European) living in Portugal (continent) and in the Atlantic Islands in different geographic and social contexts (rural, urban, proto-industrial, industrial, agrarian, administrative, commercial)
<b>4. Time period:</b> When the information of the sources was recorded? Please indicate the start and the end date.	
<b>5. Geographical area:</b> What territory is covered by the source?	Portugal (Continent and the Atlantic Islands)
<b>6. Content:</b> What was recorded?	Source, location, date of the marriage, civil status, age, occupation, residence, place of birth of the bridal pair; civil status, occupation, place of birth, residence of the parents; residence, occupation and familial relationship of the godfathers (if recorded). The completeness and administrative quality of each record varies across time and context.
<b>7. Language of written material:</b> original sources and documentation	Portuguese
<b>8. Preservation and storage:</b>	
8.a. Completely preserved	Yes
8.b. Partially destroyed by personnel according to systematic criteria	
8.c. Partially destroyed or damaged for other reasons	

8.d. Reorganized by producer of the source	
8.e. Reorganized by record linkage procedures	
8.f. Where the original records are stored (name of the archive or institution)?	In several Civil Registration Offices and District State Archives in Portugal
<b>9. Documentation:</b>	
9.a. Completely documented and accessible by:	None (so far)
9.b. Partially documented and accessible by:	
9.c. No documentation, but accessible by:	