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) and/or Tatiana Moisseenko (tatiana.moisseenko@iisg.nl)

Table of Contents:

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

1. Title of the database	BASE TRA PATRIMOINE
1.a. Subtitle, which brings meaning to the title (scope, place, time period):	
2. Abbreviation	TRA

3. Links to website(s):	
3.a. Homepage	http://tra.web.ined.fr
3.b. Get to data	http://tra.web.ined.fr/rubriques/donnees/uk_apply_data.htm

4. Abstract: describes content of the database. Max. length: 300 words Please indicate: ◦ Scope and main goal ◦ Time and territory covered by data ◦ Sample strategy ◦ Main sources	Original goal Individual and familial wealth accumulation in France (1800-1940) Time and territory covered by data The sample covers France at least a large number of departments. Sample definition: The sample is selected on the basis of last names first letters T, R, A in order to build a representative sample of Frenchmen dying from 1800 to 1940.

5. Keywords: Please use the recommended keywords if they are	Demography, civil certificates , history, occupations

applicable: *demography, life course, census, church register, civil certificates, population register, history, social science, genetics, migration, occupations.*	
Please add your own keywords, if you have data not covered by the recommended terms.	Wealth, inheritance, family reconstitution

6. Citation: Indicate how you want others to cite your database.	The TRA Project

7. IDS compatible: Indicate with *Yes* or *No* whether the database is IDS compatible, if *Yes*, please specify.	No

8. Has the database already been completed or it is still under construction?	Still under construction
8.a. If completed, please indicate the years of its construction?	
8.b. If under construction, please indicate, when it is planned to complete it?	Forever growing database
8.c. Please add a brief description of future plans for the database.	Any individual nominative source with French people or people related to France whose surname begins by the letters "TRA".

II.   Contact information

1. Name of institute or organisation	
1.a. Website	http://tra.web.ined.fr
1.b. Location: city, country	Paris, France
1.c. Postal address	PSE 48, boulevard Jourdan 75014, Paris, France
1.d. Phone	

2. Name of primary responsible person	Jérôme Bourdieu or Lionel Kesztenbaum
2.a. His/her email address	Jerome.bourdieu@ens.fr or lionel.kesztenbaum@ined.fr
2.b. Postal address	PSE 48, boulevard Jourdan 75014, Paris, France
2.c. Phone	33 1 43 13 63 67

3. Administrative information	
3.a. When this form was filled?	10/03/15
3.b. Who did it?	Jérôme Bourdieu and Lionel Kesztenbaum

4. Main economic funding (Name of organization(s) who made the grants /sustain it)	

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.

	Type of source	Yes/No	Start year	End year	Explanations:
1.	Baptisms	N			
2.	Marriages from church registers	N			
3.	Burials	N			

4.	Population registers, maintained by church or state	N			
5.	Civil birth certificates	N			
6.	Civil marriage certificates	Y	1800	1980	
7.	Civil death certificates	N			
8.	Population Census	N			
9.	Nominative lists	N			
10.	Military draft records	N			
11.	Others: wealth at death records	Y	1800	1960	
12.	Other: military registers	Y	1850	1900	

IV.   The database: core characteristics

1. Period covered by data: give first and last year of date, if possible	1800-1960
2. Territory covered by data	The sample covers France as a whole (for some part of the data it is only a large subsample of *départements*)
3. Geographical characteristic: local, regional, national, cross-national	National

4. Units of observation. Please enter *Yes* or *No* 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.

	Units of observation:	Yes /No	Number of units	Explanations:
1.	Individuals	Y		We had to assemble information on individuals that came from various locations.
2.	Married couples	Y		We have collected some information on spouse.
3.	Families	P		We have partially reconstructed family links.
4.	Households	N		
5.	Farms	N		
6.	Institutions	N		
7.	Other	N		

5. Variables per unit included in the database	
On individuals: *Data of birth and dead, age, gender, marital status, religion, occupation, migration, relationship,* etc. Please add more variables, if they are not in the list	Data of birth and death, age, gender, marital status, occupation, migration, wealth
On households: *Type of household, children present, age and number of children, etc*. Please add more variables, if they are not in the list	

6. Kinship relations:	
6.a. How is kinship recorded in the database?	We have the name of heirs. We have made partial family reconstitution.
6.b. How deep (number of generations) is kinship information going?	From two to seven generations.

7. Completeness	
7.a. Are all variables from the sources included in the database?	
7.b. Are all individuals who lived in the households of	

the sample recorded?	

8. Current data representation: Database Software (e.g. MySql, MsSql, Access, please specify	Access, STATA

9. Access conditions:	
9.a. How does a user get access to the database?	Data are available in the published book (on a CD) or can be obtained from the web.
9.b. What are the conditions and restrictions?	Restricted to 19th century data.

V.   Publications and reports

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

- Jérôme Bourdieu, Lionel Kesztenbaum , Gilles Postel-Vinay, *L'enquête TRA, histoire d'un outil, outil pour l'histoire : Tome 1 (1793-1902)*, INED, 215 p., 2014, Classiques de l'économie et de la population
- Jérôme Bourdieu, Lionel Kesztenbaum , Gilles Postel-Vinay, "The TRA Project, a Historical Matrix", *Population* 2014/2 (Vol. 69)

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

- Jérôme Bourdieu, Lionel Kesztenbaum, and Gilles Postel-Vinay (2011) "How pensions saved thrift. France at the turn of the 20th Century", *The Journal of Economic History*, 71:2 (June), p. 386-415.
- Jérôme Bourdieu, Joseph P. Ferrie, and Lionel Kesztenbaum (2009) "Vive la différence? Intergenerational Mobility in France and the United States during the Nineteenth and Twentieth Centuries", *Journal of Interdisciplinary History*, XXXIX: 4 (Spring), p. 523–557.
- Jérôme Bourdieu, Gilles Postel-Vinay, and Akiko Suwa-Eisenmann (2008), "Aging women and family wealth", *Social Science History*, 32(2), p. 143-174.
- Paul-André Rosental (1999), *Les sentiers invisibles. Espace, familles et migrations dans la France du XIXe siècle*, Paris, Eds. de l'E.H.E.S.S.
- Jacques Dupâquier and Denis Kessler (1992), *La société française au XIXe siècle, Tradition, transition, transformation*, Paris, Fayard.

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?	We observe individual at death.
2. How do individuals leave observation?	
3. How do households enter observation?	
4. How do households leave observation?	
5. Are some entry or exit dates unknown?	
6. Are some entry or exit dates estimated?	
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?	

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

1. Source(s): Which source forms the basis for the sample	Civil certificates, fiscal records
2. Sampling units: Households, individuals, regions…	Individual data
3. Variables used for selection: Age, gender, marital status, other	Selection on surname
4. Selection method: Random, stratified random, total count, clustered, other	

VIII. Data collection

1. Data collection period: When the data was collected and transcribed?	1980 -
2. Data collection method: Public digital register, transcription, other	Transcription, photo
2.a. If transcription, how was the transcription done: ◦ By individuals ◦ From scanned sources ◦ From LDS's microfilms ◦ Automatic controls	By individuals
2.b. How was the checking of the transcription done? For example, by proof reading?	
2.c. When was it done?	
2.d. Purpose of the transcription: please indicate ◦ LDS ◦ Research ◦ Genealogy	research
3. Control methods by researcher: e.g. Internal consistencies such as a death cannot happen before a birth of the same person	Internal consistency
4. Data collection staff: Please indicate the number of people and their position (member of the project, free-lancer, other)	3 full time researchers, ½ time not permanent assistants

IX.   Linkage process

1. Linkage: Which sources and units of observation have been linked: (e.g. birth/baptisms and death/burials…)?	Marriage and fiscal sources
2. Documentation of linking:	
2.a. Programme, manually, …	Both programme and manual
2.b. Name of software if used (and its parameters)	
3. What are the rules for linking? Flags definition (list them: age, name, extra knowledge …)	
4. How each reconstructed person is traceable to the original sources /transcribed data?	
5. How is linkage represented in the database? For example, do all occurrences of an individual include a universal identification number (ID)? Or are records linked in another way?	
6. Linkage percentage	50%
7. Quality of linkage (own evaluation)	80%
8. What reference/coding systems have been linked to the data? 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).	

Y/N/P	*Reference system*	*Explanations:*
Y	Occupational titles:	Occupational titles (like HISCO).
Y	Locations (including geo-referenced systems):	Locations (including geo-referenced systems).
	Religion, civil status etc.:	
	Other:	

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)

1. Official name of the source and its English translation	
2. Purpose of the source:	
2.a. Why was this source created?	
2.b. Who created it?	
3. Scope: What group of the population was documented in this source?	
4. Time period: When the information of the sources was recorded? Please indicate the start and the end date.	
5. Geographical area: What territory is covered by the source?	
6. Content: What was recorded?	
7. Language of written material: original sources and documentation	
8. Preservation and storage:	
8.a. Completely preserved	
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)?	
9. Documentation:	
9.a. Completely documented and accessible by:	
9.b. Partially documented and accessible by:	
9.c. No documentation, but accessible by:	