



# THE HHB DATABASE CODEBOOK

---

DO NOT CITE, CIRCULATE, OR COPY WITHOUT PERMISSION

Draft

<http://hhbproject.com/>

[info@hhbproject.com](mailto:info@hhbproject.com)

## Introduction

---

Should one use data if one does not know how they are constructed?

The aim of the **Historical Household Budgets (HHB)** project is to rewrite the history of both inequality and poverty over the last two centuries, and to make that history international. Existing accounts perch on a narrow evidentiary base: few countries, few indicators (often proxies), short time frames, poor comparability across sources and countries, unrepresentative groups such as the top 1 percent. A hitherto underutilized source of data, household budgets, can provide the foundation for better estimates of long-run changes in income distribution.

The initial focus of the HHB project is the construction of a multi-topic database collecting household budgets from 1850 to the present day. We call it the **Historical Household Budgets Database (HHBD)**.

For the purposes of the HHBD, a **household budget** comprises at minimum the following five pieces of information:

- i)* time – the year to which the data refer;
- ii)* place – the family's region and residential status;
- iii)* size – the number of individuals making up the household;
- iv)* occupation – the head of household's profession and/or sector of activity;
- v)* household resources – either total income or total expenditure.

While the list above provides the list of essential ingredients, many historical budgets are richer in their information.

The HHBD comprises ca. **500 variables** including monetary measures of the standard of living (consumption, income, wealth, but also wages and retail prices), education and health, anthropometric measures, fertility, employment and migration, housing, agriculture, access to credit, and exposure to shocks.

In addition to household budgets, the HHBD contains an extensive collection of [digitalized references](#). These are mainly the sources from which household budgets are extracted.

### [Sources]

This **Codebook** contains a detailed description of the variables constituting the database. (...)

A companion **Manual** addresses the issues that arise at the data-entry level, and is meant to assist HHBD users who engage the data entry phase.

## Structure of the database

The design of the HHBD is based on a number of sources. A first key reference is the standard set by the World Bank's **Living Standard Measurement Study** (Grosh and Glewwe, 2000). Its design has influenced (and still is influencing) most modern household budget surveys in low- and middle income countries all around the world. A second key reference is the **Luxembourg Income Study's** (LIS) documentation. The LIS Project (...) efforts to harmonize micro-datasets from upper- and middle-income countries (Smeeding, Schmaus, and Allegrezza, 1985). Many other references have been used, and will be mentioned as required along the way of this Codebook.

The HHBD is organized in **18 sections**, which contain information at the household, individual and item level, as well as information about the original source of the data.

Information level – Legend					
<i>H</i>	Household level	<i>A</i>	Individual level	<i>t</i>	Item level

### Sections and contents

#### 1. Metadata *H*

*This section collects basic information on the source, as well as on defining features of the budget, such as the reference year, the currency unit and much more.*

#### 2. Household Roster *HA*

*Selected socio-demographic variables at the household and individual level.*

#### 3. Health, Fertility and Anthropometrics *At*

*Health status of household members, access to health services, as well as anthropometric indicators and fertility.*

#### 4. Education *A*

*Educational achievements and school participation.*

#### 5. Migration *A*

*Internal and international migration history.*

#### 6. Expenditure *t*

*Household expenditures, by item. Also taxes are reported here.*

<b>7. Durables</b>	t
<i>Durable goods owned by the household.</i>	
<b>8. Wealth</b>	t
<i>Household wealth by asset type.</i>	
<b>9. Housing</b>	H
<i>Housing conditions of the household.</i>	
<b>10. Credit</b>	t
<i>Access to credit and financial services.</i>	
<b>11. Income</b>	t
<i>Household sources of income.</i>	
<b>12. Employment</b>	At
<i>Employment status, experience and skills, by individual; working conditions and job characteristics (sector, duration, wage, etc.), by job; task details, by task, when available.</i>	
<b>13. Agriculture</b>	t
<i>Agricultural activities of household members in detail, by activity.</i>	
<b>14. Non-Agricultural Enterprise</b>	t
<i>Non-agricultural household enterprises, by activity.</i>	
<b>15. Time Use</b>	t
<i>Allocation of time of the household members.</i>	
<b>16. Shocks</b>	t
<i>Unexpected events impacting on the household in the past year.</i>	
<b>17. Community Data</b>	t
<i>Access to facilities available in the community.</i>	

The Codebook devotes a chapter to each module (...).