



NTNU  
Norwegian University of  
Science and Technology

BANCO DE ESPAÑA  
Eurosistema



VNIVERSITAT DE VALÈNCIA [0%] Facultat  
d'Economia



# Two stories, one fate:

## Age-heaping and literacy in Spain 1877-1930

Francisco J. Beltrán Tapia

*Norwegian University of  
Science and Technology*

Alfonso Díez Minguela

*Universitat de València*

Julio Martínez Galarraga

*Universitat de València*

Daniel A. Tirado Fabregat

*Universitat de València*

# 1. Introduction

- Human capital and economic development (Romer, 1986; Lucas, 1988; Gennaioli *et al.*, 2013).
- Did **human capital** have an effect on (long-run) **regional development** in Spain?
- Several obstacles, above all, **measurement**, e.g. what is human capital?
  - Reading and writing skills
  - Numeracy skills
  - Other skills

# 1. Introduction

- As regards Spain, the existing literature (i) is mostly qualitative, except for Núñez (1992), or (ii) explores region-specific cases.
- This study examines **human capital** in the early stages of modern economic growth in Spain. In doing so, we use all the population censuses since 1860.
- Interestingly, this exercise *“raises more questions than answers”*.

## 2. Methodology and data

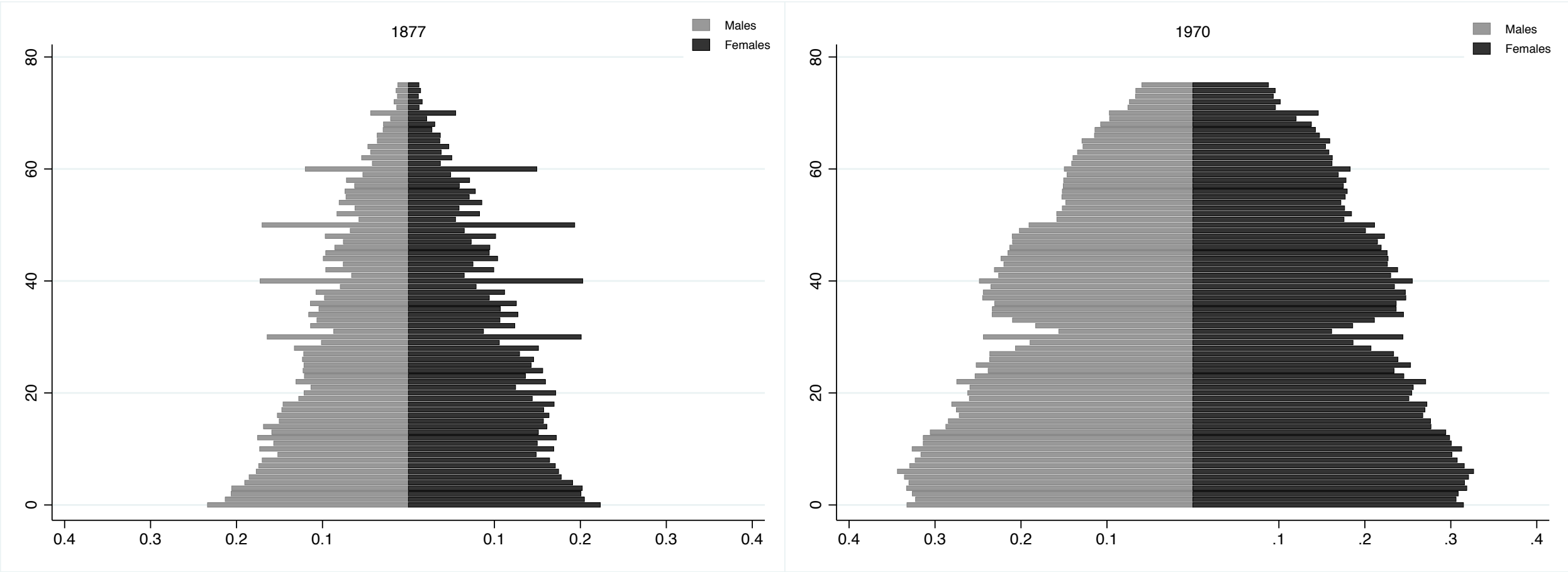
- Proxies of human capital? (literacy; enrolment...)

## 2. Methodology and data

- Proxies of human capital? (literacy; enrolment...)
- **Age-misreporting** is a common phenomenon in historical records, e.g. civil and legal documents, military records, passenger lists, parish registers or censuses.

# 2. Methodology and data

**Figure.** Population pyramid by age (0-75 years old) and gender in 1877 and 1970



Source: INE.

## 2. Methodology and data

- Proxies of human capital? (literacy; enrolment...)
- **Age-misreporting** is a common phenomenon in historical records, e.g. civil and legal documents, military records, passenger lists, parish registers or censuses.
- In economic and social history, **age-heaping** usually related to poor numeracy skills (A'Hearn et al., 2009; Crayen & Baten, 2010), custom and/or tradition.

## 2. Methodology and data

- Several tools to gauge age-heaping (Bachi's index, Myers' index, **Whipple index**, Zelnik's index...).
- The **Whipple index** assumes that respondents are uniformly distributed over an age-range (23-62; 23-72...):

$$W = \frac{(P_{25} + P_{30} + P_{35} + P_{40} + P_{45} + P_{50} + P_{55} + P_{60})}{1/5 (P_{23} + P_{24} \dots + P_{61} + P_{62})}$$



## 2. Methodology and data

- Then, the Whipple index is transformed into an **ABCC index (%)**:

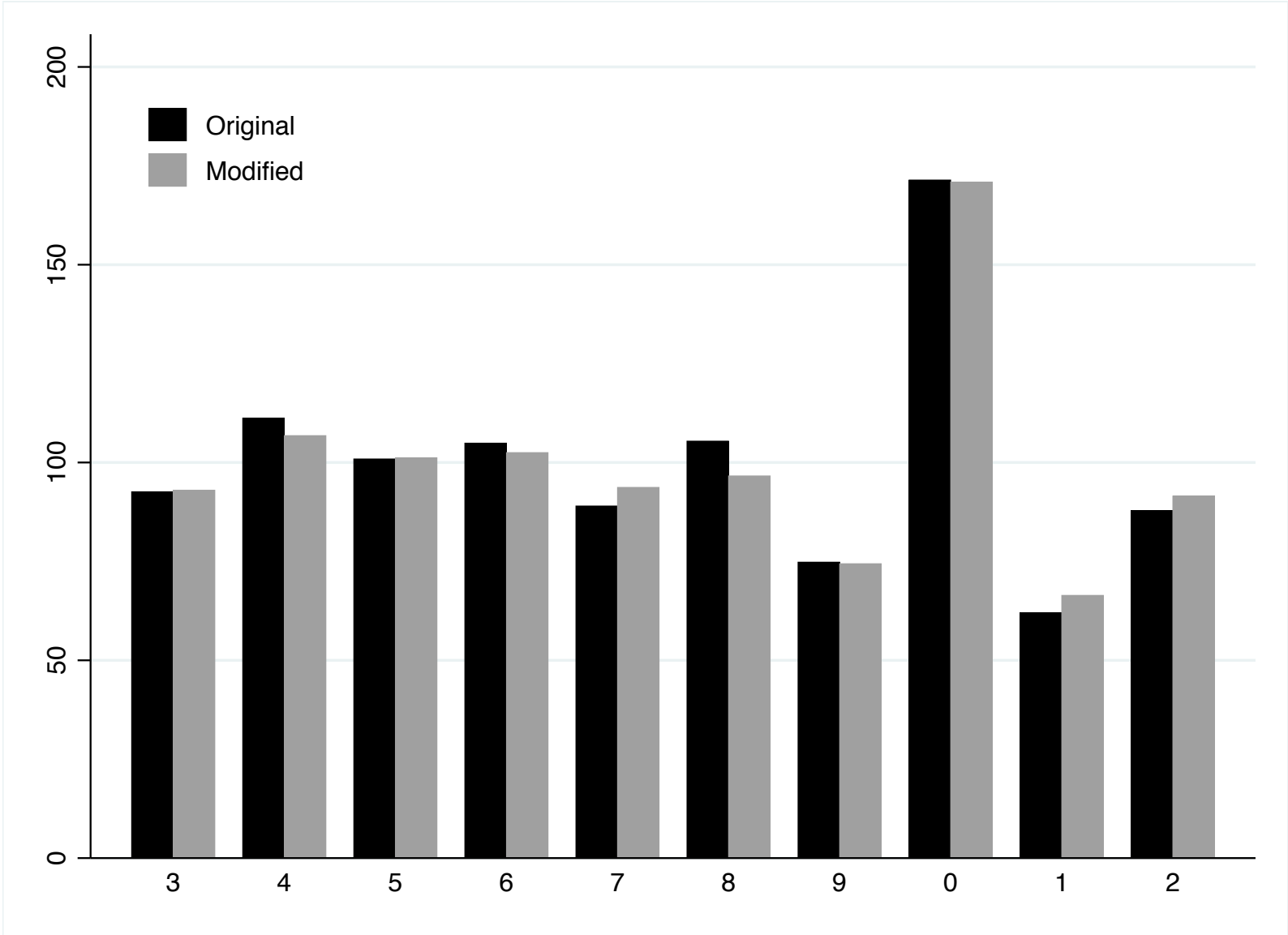
$$ABCC = \left\{ 1 - \frac{(W - 1)}{4} \right\} \times 100 \text{ for } W \geq 1$$

$$ABCC = 100 \text{ elsewhere}$$

- But, **Spanish censuses** do NOT exhibit a preference for ages ending in 5.
- What then? Noubissi (1992) proposed a **modified Whipple index** to account for preference/avoidance of all digits.

# 2. Methodology and data

**Figure.** Digit preference (23-62 years) by method in the census of 1877.



Source: INE.

### 3. Descriptive analysis

- Why doing this? A first glance at Spain in mid-nineteenth century.

### 3. Descriptive analysis

**Table.** Age-heaping in Europe during the 1850s.

Country	ABCC index (%)
Belgium	100.0
Finland	100.0
France	100.0
Sweden	100.0
Switzerland	100.0
Germany	99.7
Italy	99.4
Denmark	99.2
Netherlands	99.0
Norway	98.4
Austria	98.0
UK	97.1
Spain	87.4
Russia	83.6

**Note:** ABCC index (%) illustrated above is a birth decadal average.

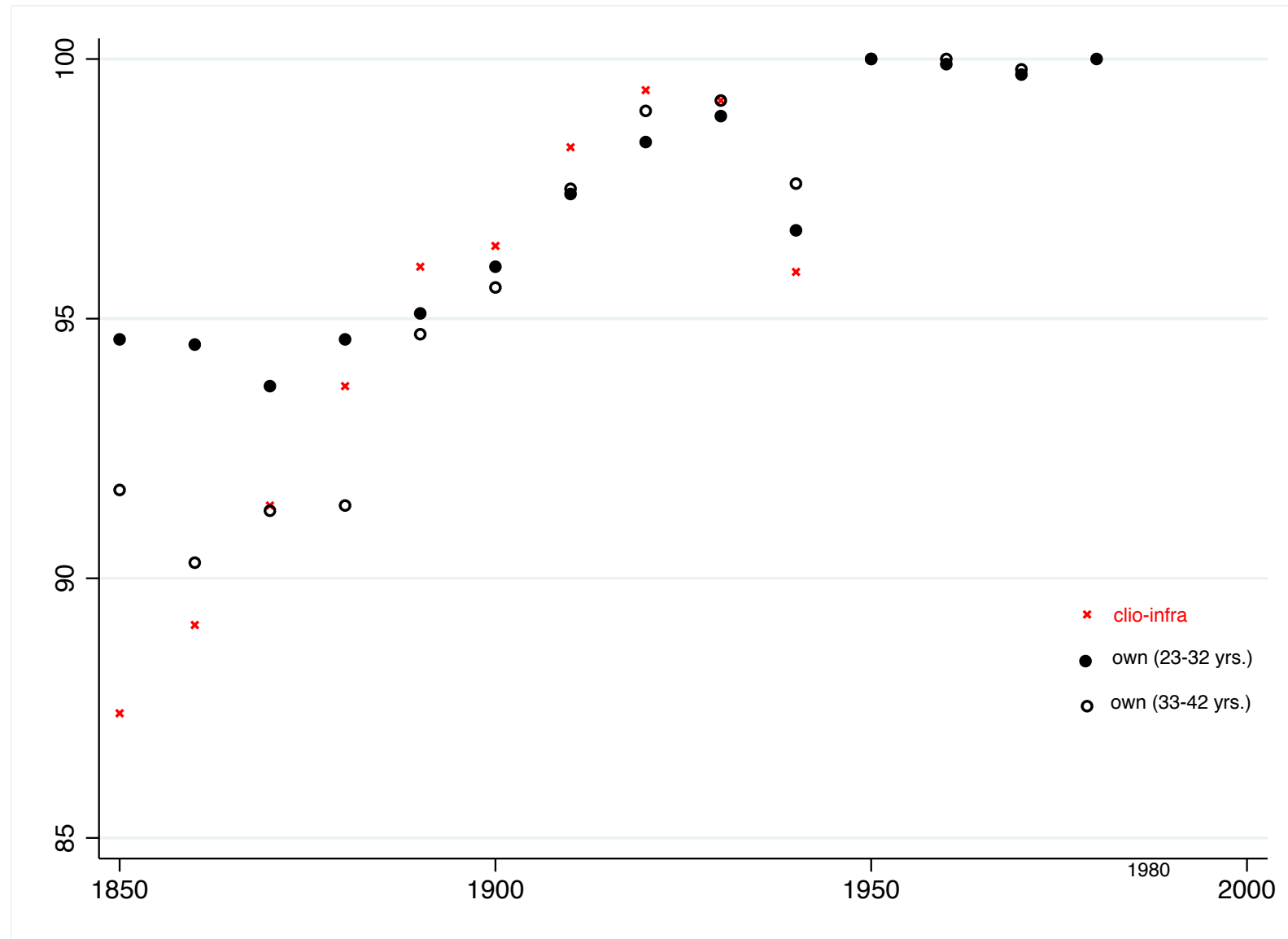
**Source:** clio-infra (<https://www.clio-infra.eu>)

### 3. Descriptive analysis

- Why doing this? A first glance at Spain in mid-nineteenth century.
- Still, the **clio-infra database** used those aged 43-52 yrs. in the 1900 census. If, for instance, we use information for 23-32 or 33-42 in previous censuses, then...

### 3. Descriptive analysis

**Figure.** ABCC index (%) in Spain by birth decade, 1850-1980.



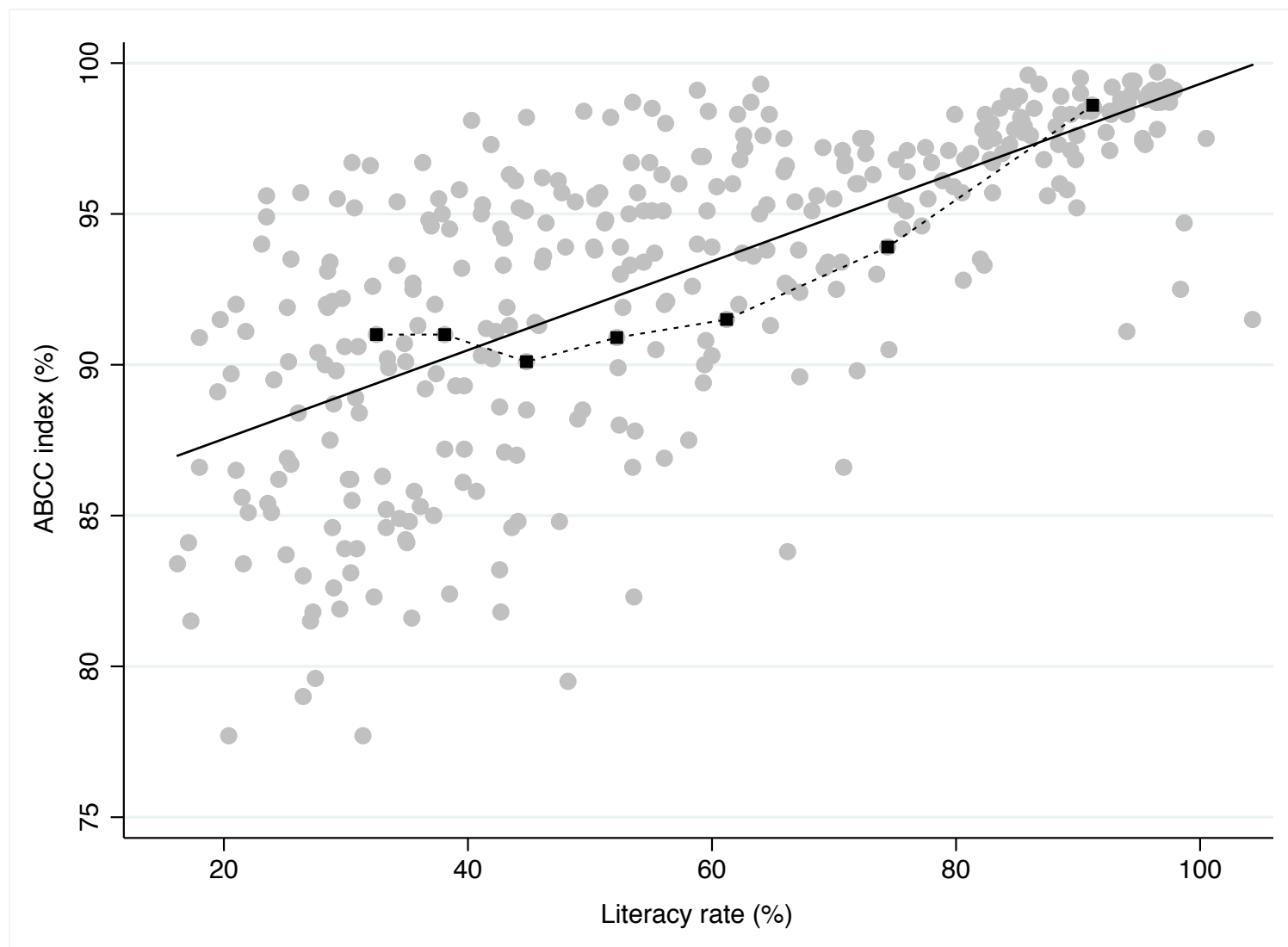
Source: Clio-infra; INE.

### 3. Descriptive analysis

- Why doing this? A first glance at Spain in mid-nineteenth century.
- Still, the **clio-infra database** used those aged 43-52 yrs. in the 1900 census. If, for instance, we use information for 23-32 or 33-42 in previous censuses, then...
- Interestingly, **age-heaping** did not improve in late nineteenth-century...but, what about **literacy**?

### 3. Descriptive analysis

**Figure.** ABCC index (%) and literacy rate in Spain by province for a selection of censuses.



**Notes:** Data for the following population censuses: 1877, 1887, 1900, 1910, 1920, 1930 and 1970.

**Source:** Núñez (1992); INE and authors' calculations.



### 3. Descriptive analysis

- Then, are age-heaping and literacy comparable?

ABCC index (% of individuals reporting age correctly, 23-62 yrs.)

Literacy rates (% of literates, 10 or more yrs.)

- Though age-heaping and literacy are correlated, this relationship not as strong as in other studies (A'Hearn et al., 2009).

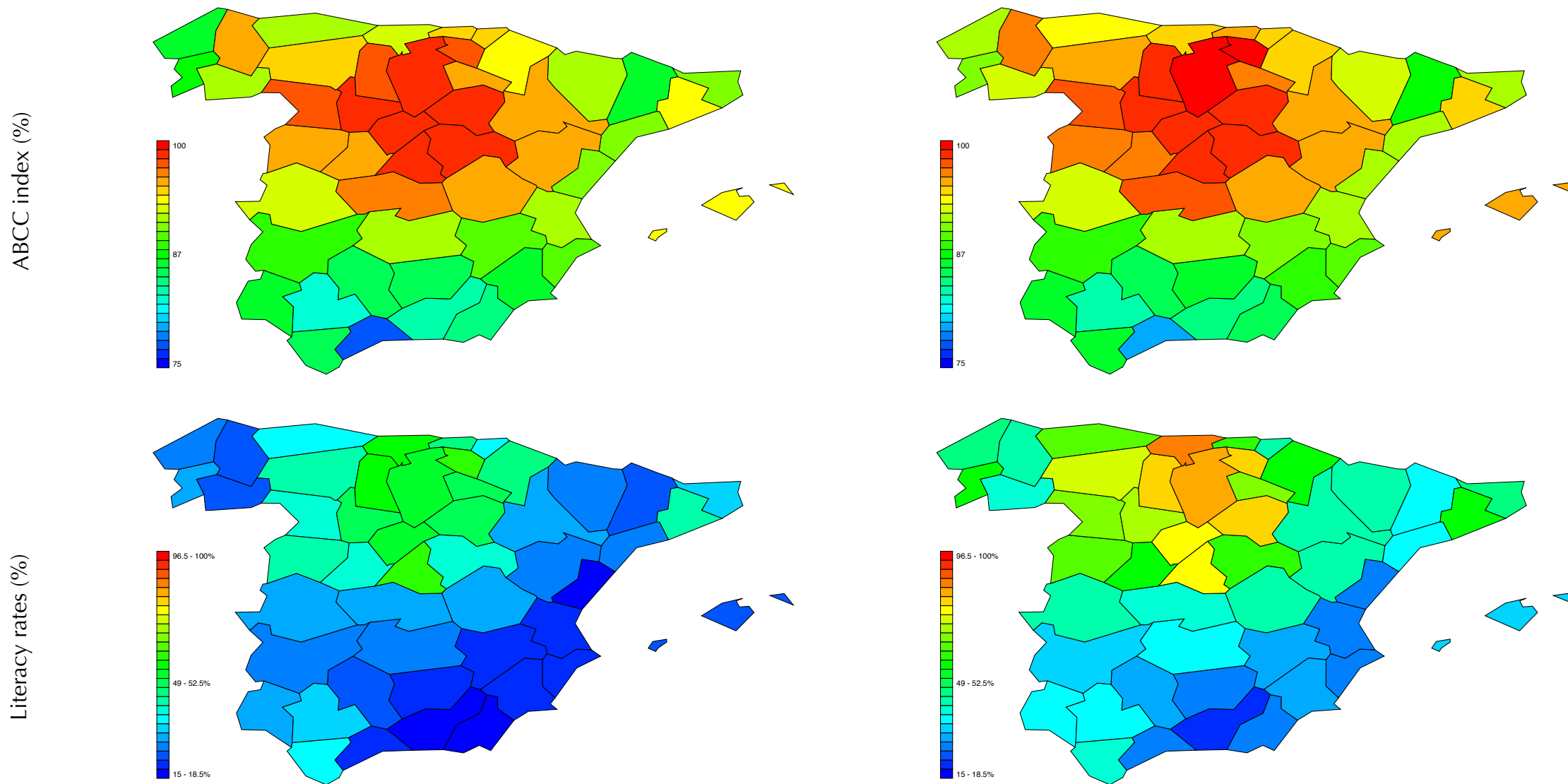
- Furthermore, was not age-heaping “*surprisingly*” low?

ABCC index in Burgos, Guadalajara, Madrid, Segovia, Soria, Valladolid above 95%

Literary rates in Burgos, Guadalajara, Madrid, Segovia, Soria, Valladolid ranging from 40.3% to 62.1%.

### 3. Descriptive analysis

**Figure.** Age-heaping and literacy in Spain, 1877. (Total-left; Male-right)



**Notes:** Provincial ABCC index (%) and literacy rates (%) classified into 25 equivalent categories for all censuses ranging from the minimum value (dark-blue) to the maximum or 100 (dark-red).  
**Source:** INE and authors' calculations.

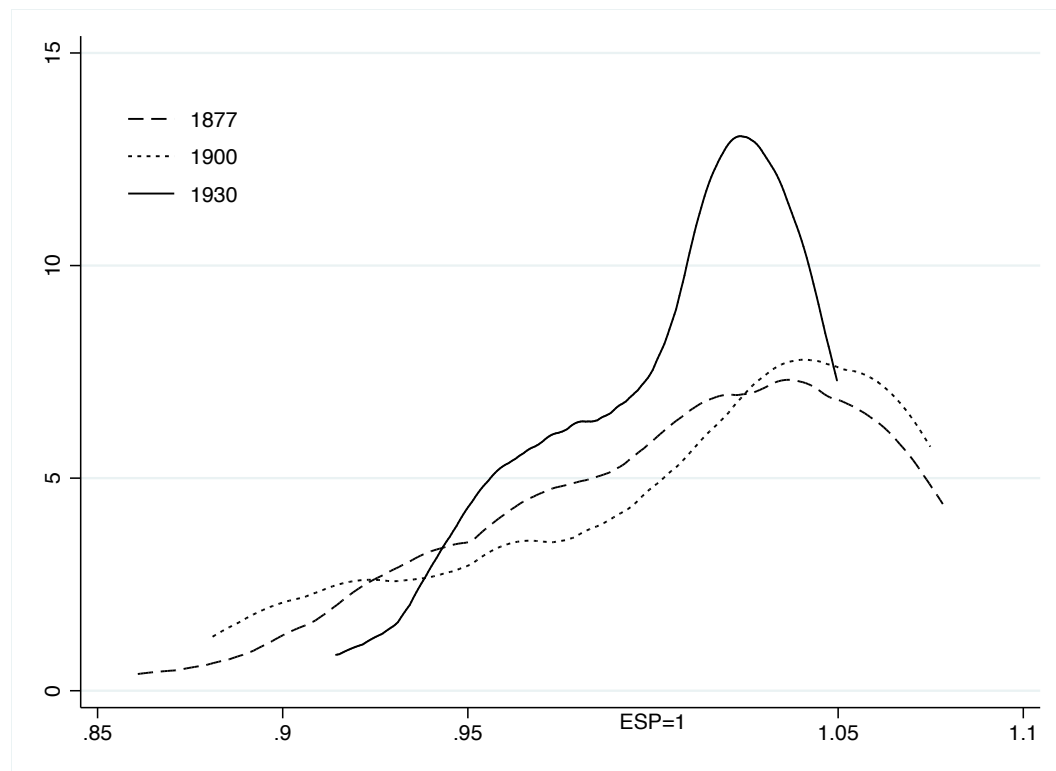
### 3. Descriptive analysis

- Even more, **gender differences** in age-heaping are trivial, but this is not the case for literacy..."self-reported" age?
- Also...the dynamics of age-heaping and literacy (only male data) differed widely.

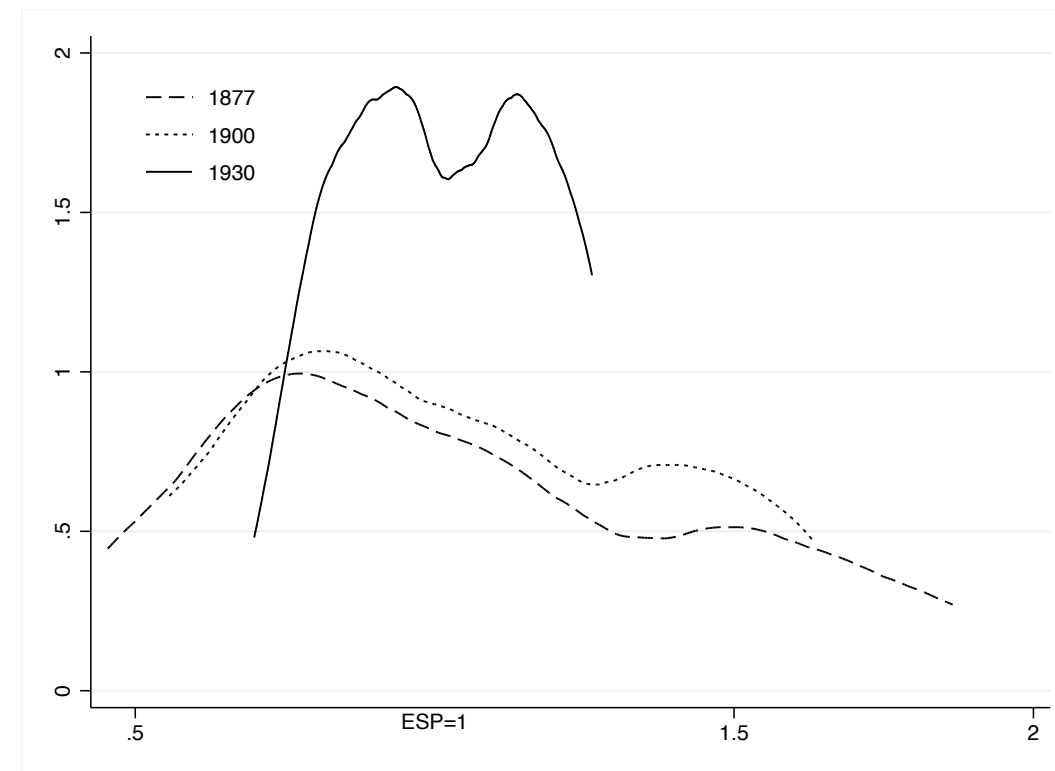
### 3. Descriptive analysis

**Figure.** Male age-heaping and literacy kernel densities in Spain in 1877, 1900, and 1930.

(a) ABCC index (%) (ESP=1)



(b) Literacy rates (%) (ESP=1)



**Notes:** Provincial ABCC index (%) and literacy rates (%) ranging from the minimum value to the maximum or 100.

**Source:** INE and authors' calculations.

### 3. Descriptive analysis

- Even more, **gender differences** in age-heaping are trivial, but this is not the case for literacy..."self-reported" age?
- Also...the dynamics of age-heaping and literacy (only male data) differed widely.
- Are age-heaping and literacy telling us two distinct stories?

## 4. Conclusion

- Our main findings can be summarised as follows:
  - #1# Spanish censuses exhibit a preference for digits ending in 0.
  - #2# In Spain, age-heaping did not improve until early twentieth-century (Note: Beware of age-effect)
  - #3# Given the level of literacy, age-heaping appears to be abnormally low, especially in some provinces.
  - #4# There seems to be a minor and insignificant gender gap in age-heaping.
- The study thus “casts doubt on the extent to which **digit preference** proxies **numeracy skills**, and henceforth human capital”, as A’Hearn *et al.*, (2016) did for Italy.
- As pointed above, the study “raises more questions than answers”, especially regarding the efficacy of the public administration, hence calling for **further research**.