Census and Administrative Data in Portugal: Results and Challenges

Joana Araújo, Cecília Cardoso, Paula Paulino, Cláudia Pina, Pedro Raposo

Statistics Portugal

Census Unit

INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL

eurostat

June 5th, 2024

1

Outline

EUROPEAN CONFERENCE ON QUALITY IN OFFICIAL STATISTICS 2024 ESTORIL - PORTUGAL

INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL

eurostat

The conference is partly financed by the European Union

2

1. From Traditional to Administrative Census

▷ Statistics Portugal is preparing to change the Census model

▷ The administrative Census project is part of the development of a National Data Infrastructure (IND) which embodies Statistics Portugal's strategy of integration and creation of value for society from different data sources

▷ The administrative Census challenges:

 ▷ There is no administrative population register

 ▷ There is no unique Personal Identification Number (PIN)

 ▷ Linking administrative data

EUROPEAN CONFERENCE ON
QUALITY IN OFFICIAL STATISTICS
2024 ESTORIL - PORTUGAL

INSTITUTO NACIONAL DE ESTATÍSTICA
Statistics Portugal

eurostat

The conference is partly
financed by the European Union

1. From Traditional to Administrative Census

▶ The central element of the Administrative Census project is the Resident Population Database (**BPR**) - a statistical database, covering a set of characteristics (geographical, demographic and socio-economic), that results from the integration of various sources of the public administration

▶ The database of all Portuguese nationals (Civil Identification Database) and a database of non-nationals with a valid resident permit or an EU resident card are key elements of **BPR**

▶ To establish the usual residence in Portugal, "signs of life" rules are applied. The "signs of life" are given by the presence of the person in other administrative register (Income Statement, Social Security, Education, …)

4

2. Comparison between Census and BPR
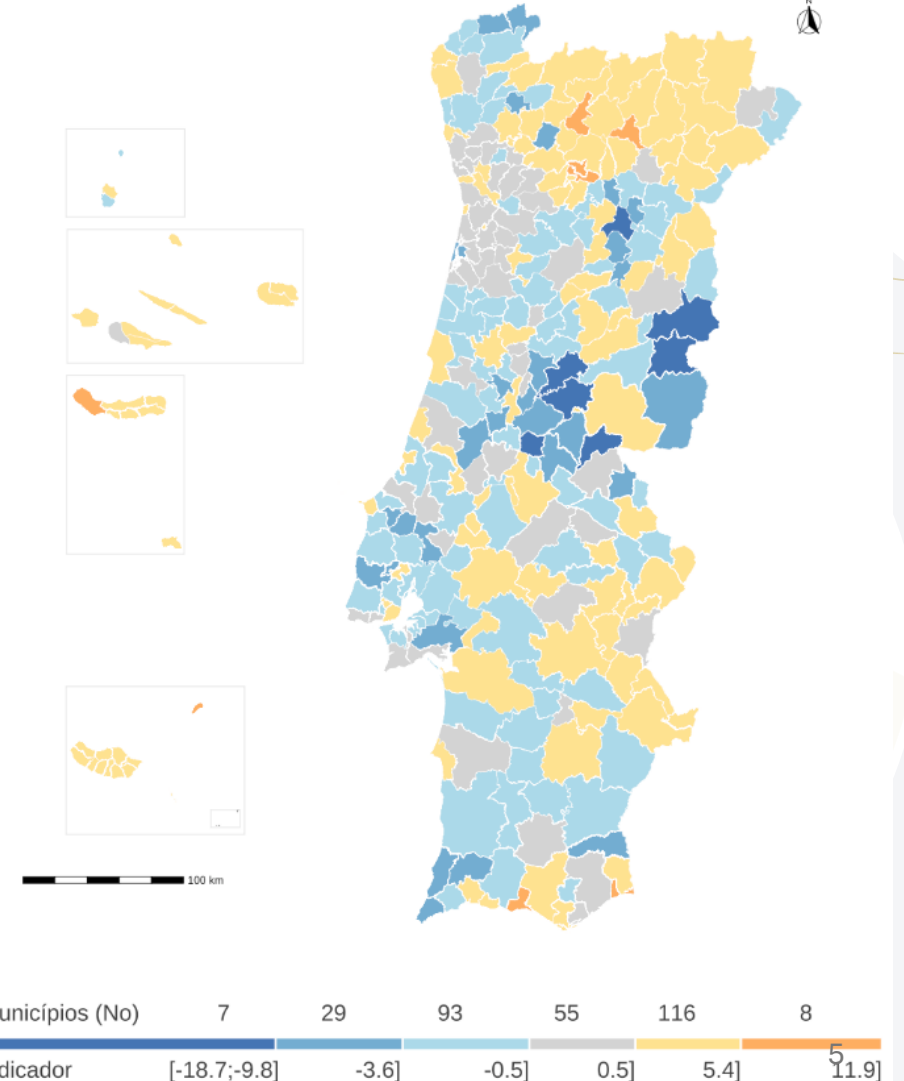## 2.1 Aggregated results

Population difference (%) between Census and BPR, by NUTS2

NUTS 2	BPR [1]	CENSOS [2]	Difference (%)
Portugal	**10,384,936**	**10,369,900**	**0.1**
Norte	3,604,970	3,593,339	0.3
Centro	2,220,072	2,235,372	-0.7
AM Lisboa	2,866,336	2,877,213	-0.4
Alentejo	709,280	707,791	0.2
Algarve	477,507	468,335	2.0
R. A. Açores	247,210	236,620	4.5
R.A. Madeira	259,561	251,230	3.3

[1] BPR reference date: December 31st, 2020

[2] Census reference date: April 19th, 2021, adjusted to December 31st, 2020

Population difference (%) between Census and BPR, by Municipality

| Municípios (No)	7	29	93	55	116	8
Indicador	[-18.7;-9.8]	-3.6]	-0.5]	0.5]	5.4]	11.9]

2. Comparison between Census and BPR
2.1 Aggregated results

- Similar age structure

- Overestimation in the male population aged between 23 and 49

- Underestimation in female population between the ages of 40 and 67

- Overestimation in the elderly population

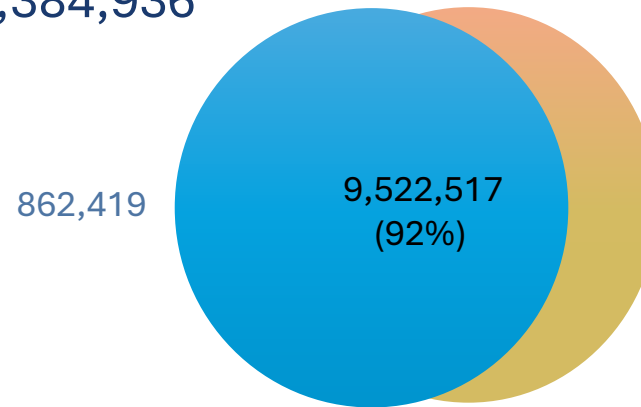**Age pyramid, BPR and Census (% of total population), Portugal**

MALES

FEMALES

☐ BPR
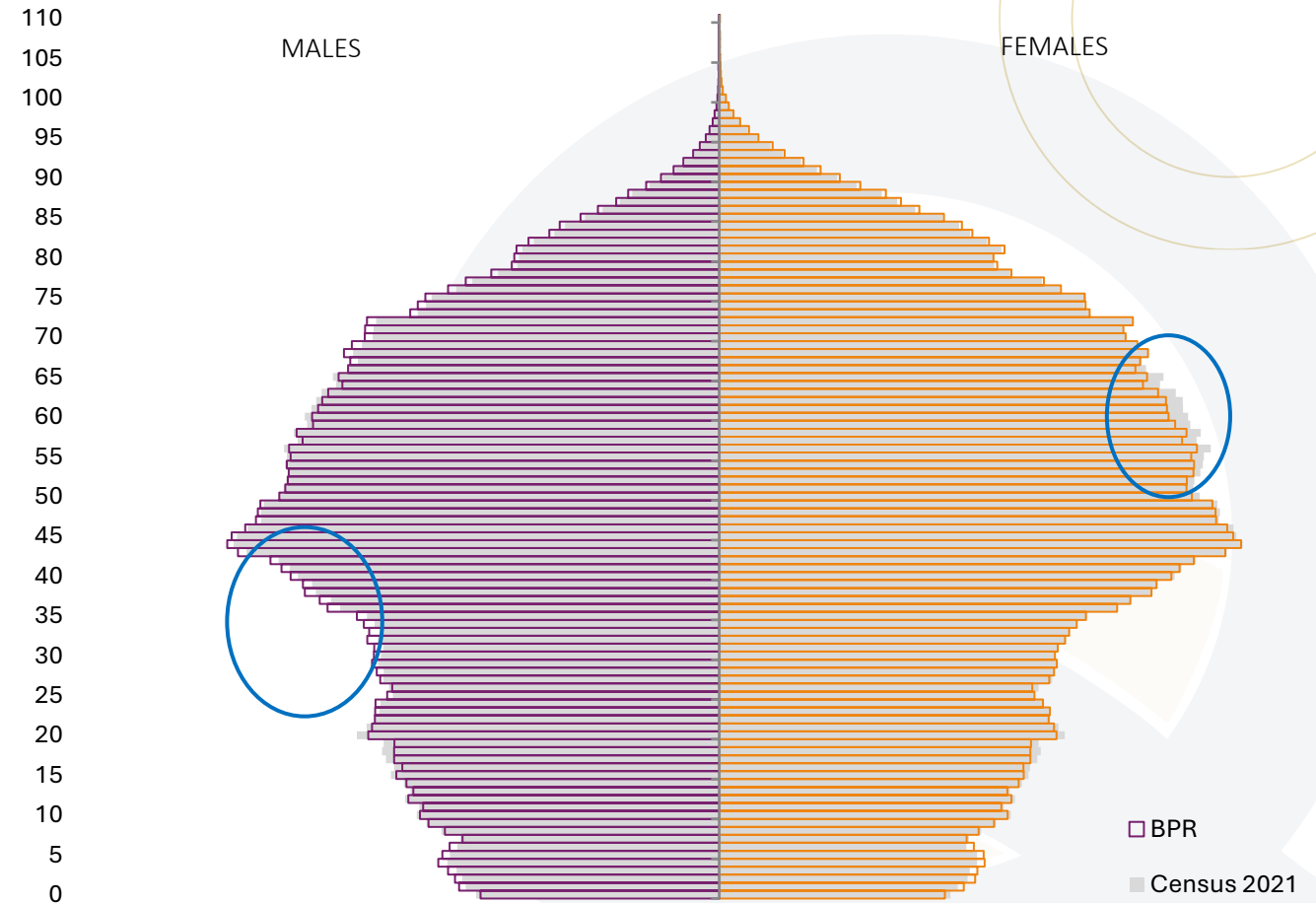
■ Census 2021

2. Comparison between Census and BPR
2.2 Micro data level: Matched individuals

2020 BPR: 10,384,936

2021 Census: 10,343,066

862,419

9,522,517
(92%)

820,549

Techniques used in linkage (name and birth date) between BPR and Census:

▷ Deterministic

▷ Similarity distance (Jaro Winkler and Levenshtein)

EUROPEAN CONFERENCE ON
QUALITY IN OFFICIAL STATISTICS
2024 ESTORIL - PORTUGAL

INSTITUTO NACIONAL DE ESTATÍSTICA
Statistics Portugal

eurostat

The conference is partly
financed by the European Union

2. Comparison between Census and BPR
2.2 Micro data level: Matched individuals

Similarity rates between in Census and BPR (%)

Category	Value
Sex	100.0
Citizenhip	99.5
Age	99.4
Place of birth	99.0
Marital status	97.5
Place of residence (Municipality)	95.4
Place of residence (LAU/parish)	91.8
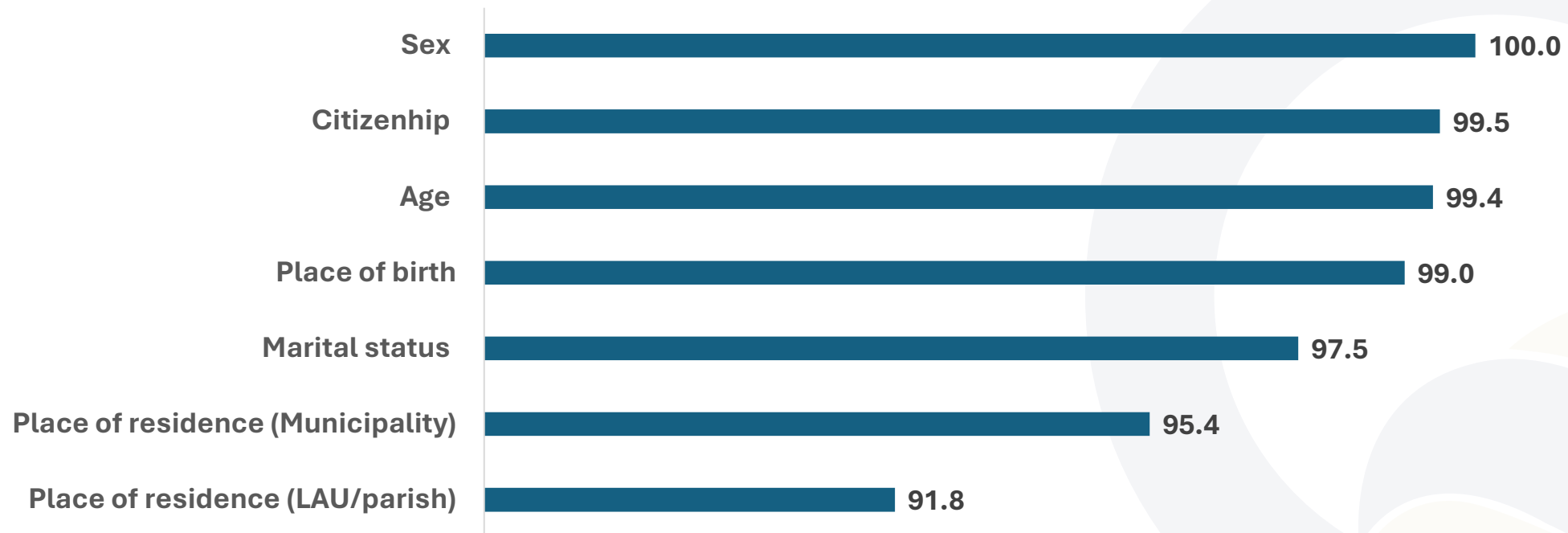
2. Comparison between Census and BPR
2.2 Micro data level: Matched individuals

Probit model (marginal effects) to understand the characteristics of the individuals whose LAU/parish of residence is different in the two datasets

- Asian (15.5 p.p) and African (10.5 p.p) nationals

- Aged above 85 (7.4 p.p)

- Divorced (6.2 p.p)

	Marginal Effect
Sex (vs Male)	
Female	0.006
Age (vs]40-85])	
[0-5]	0.041
]5-20]	0.021
]20-40]	0.046
>85	0.074
Marital status (vs married)	
Divorced	0.062
Single	0.046
Widowed	0.037
Citizenship (vs Portuguese)	
America	0.098
Oceania	0.098
Other countries in Europe	0.039
European Union	0.045
Africa	0.105
Asia	0.155

2. Comparison between Census and BPR
2.3 Micro data level: Unmatched individuals

▷ The analysis of the records found in census and not in the BPR (820,549) was also decisive to understanding the effectiveness of the BPR's methodological approach, particularly regarding the "signs of life" rules

▷ Excluding the records that could not be matched due to lack of linking attributes and the fact both databases had different reference dates (e.g. children born after the December 31st, 2020 can not be found in 2020 BPR), it was possible to identify a set of individuals in some administrative sources, but which were not integrated into the BPR (restrictive rules/ inconsistent individuals record data between sources)

▷ Non-Portuguese population and some specific groups of non-economically active population were more represented in these unmatched records

3. Conclusions and takeaways

- High level of convergence between the results obtained from the Census (traditional model) and BPR (administrative data) in several dimensions

- 92% of match individual records and similarity rates above 97% for demographic variables

- Non-Portuguese population, older population and divorced population have higher probability of having a different place of residence in Census and in BPR

- Need to improve the coverage of some groups (non-Portuguese population, specific groups of non-economically active population) through the review of "signs of life" rules

- Access to new administrative data sources will improve coverage and social-economic characteristics of the individuals

- The results of this study allow us to anticipate a successful transition to an administrative census model

EUROPEAN CONFERENCE ON
QUALITY IN OFFICIAL STATISTICS
2024 ESTORIL - PORTUGAL

Thank you

joana.araujo@ine.pt

INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL

eurostat

The conference is partly
financed by the European Union