

Fertility of Emigrants via Online Retail Data: The Case of MammaPack

Francesco Rampazzo, Jason Bell, Ridhi Kashyap, Micol Morellini, and Doug Leasure, Melinda Mills, and Andrew Stephen

Research Question

Is it possible to use **consumer data** 🛒 sources to comprehend the **connection** between **migration** 🧳 and **fertility** 👶 in both sending and receiving countries?

Background

Migration and fertility are intertwined phenomena presenting challenges for researchers due to data limitations, impeding our understanding of their impact on population dynamics.

Data

The primary data source is **MammaPack** 📦, an online retail platform founded in 2015. MammaPack serves **21 European countries**, boasting **20,000+ customers** and **10,000+ products** 🍷 from 900 brands. Our dataset covers transactions from **January 2019 to December 2022**, including customer postcodes, purchased items, and dates. We conducted an **online survey** of MammaPack customers between June and July 2022, receiving responses from **700 participants** and linking them to their purchase histories.

Hypothesis

- MammaPack customers are expected to be **concentrated** in **specific regions** of Europe 🌍, particularly in identified **“hot spots”** based on the geographic distribution of migrants (Lichter et al., 2020).
- Our hypothesis is that Italian migrants, as suggested by Anelli and Balbo (2021), may opt to **have children abroad** 👶 rather than in Italy, potentially influenced by fertility choices and different regulatory environments (Mussino and Cantalini, 2024). We predict **higher fertility rates** in these **regions** compared to rates **in Italy**, which we will confirm through the survey.

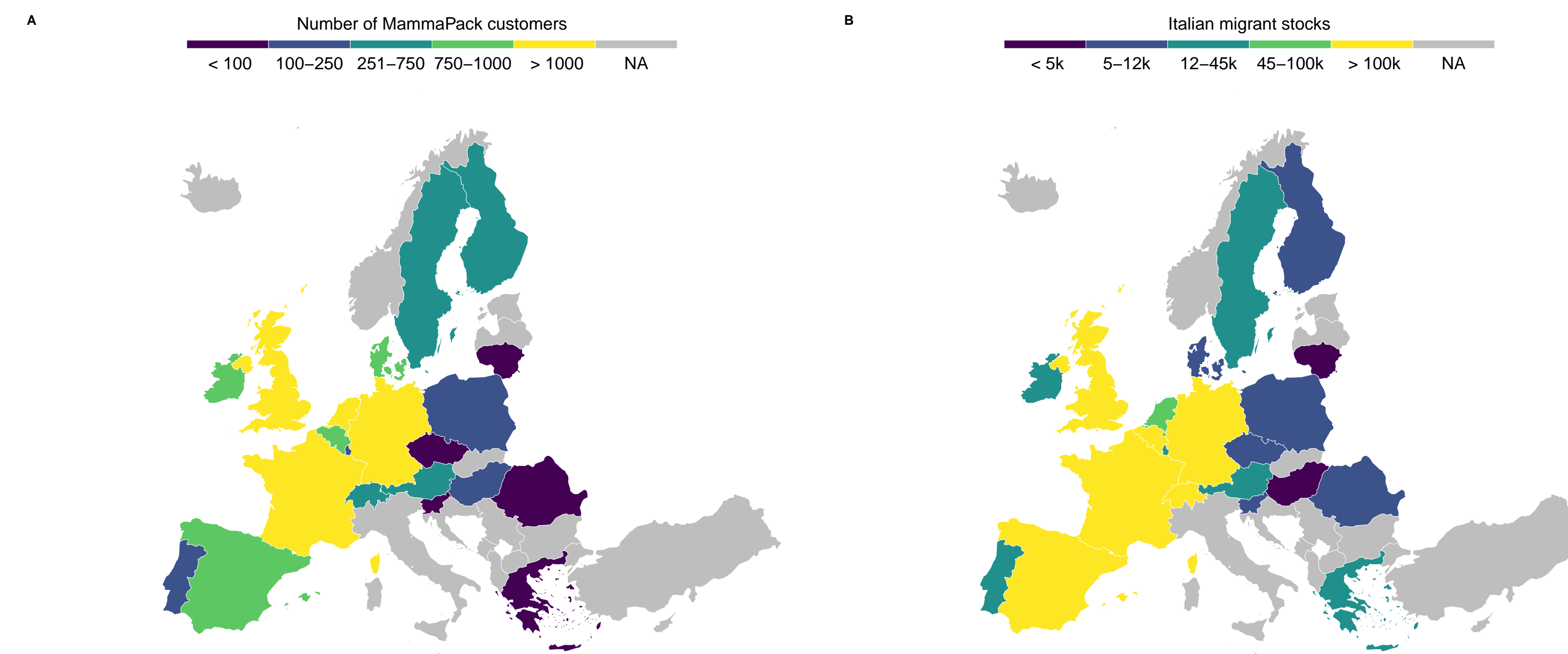


Figure 1: Maps showing (A) the number of unique MammaPack customers as of January 2023; and (B) the stock of Italian migrants registered by the Italian Registry of Italians Resident Abroad (AIRE) in 2021, both at the national level.

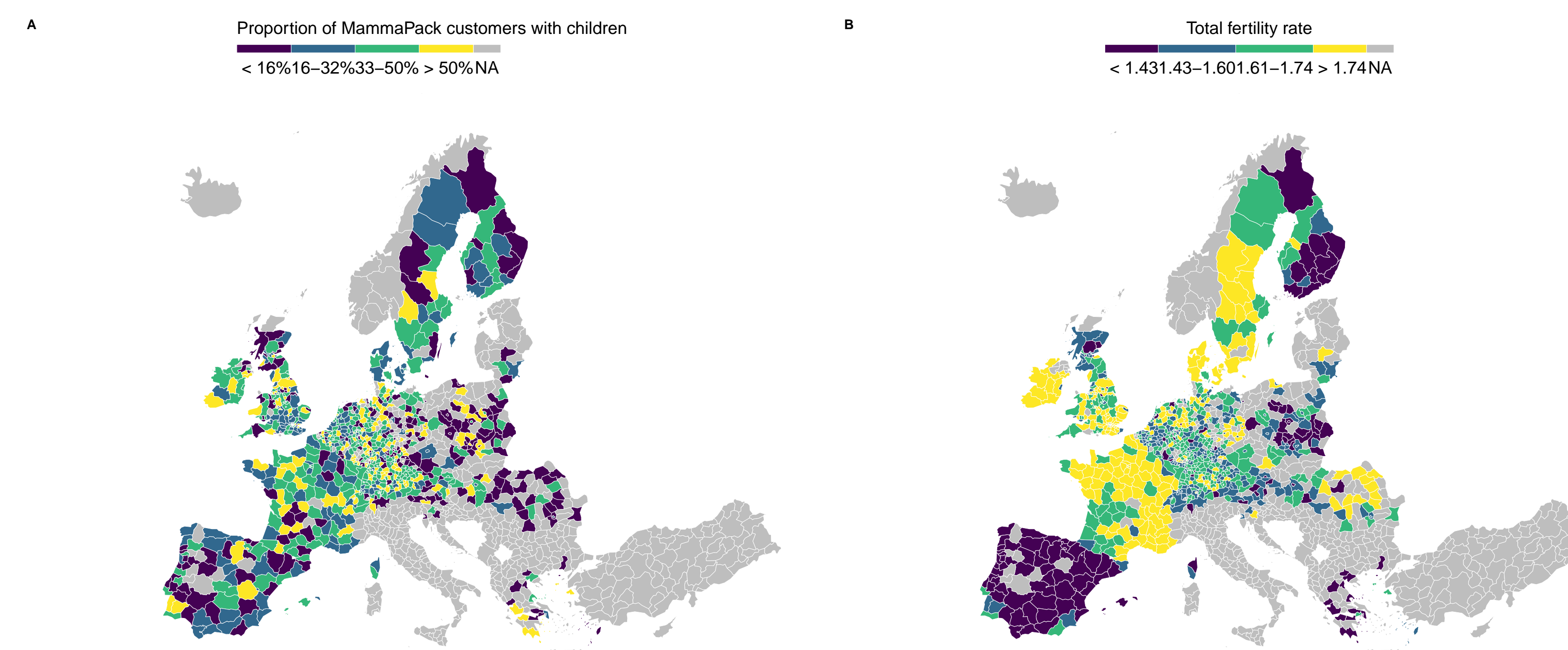


Figure 2: Maps depicting (A) MammaPack customers predicted to have a child during the period 2019-2022; and (B) total fertility rates in corresponding NUTS3 regions in 2018. Both measures are colour-coded by quartiles.

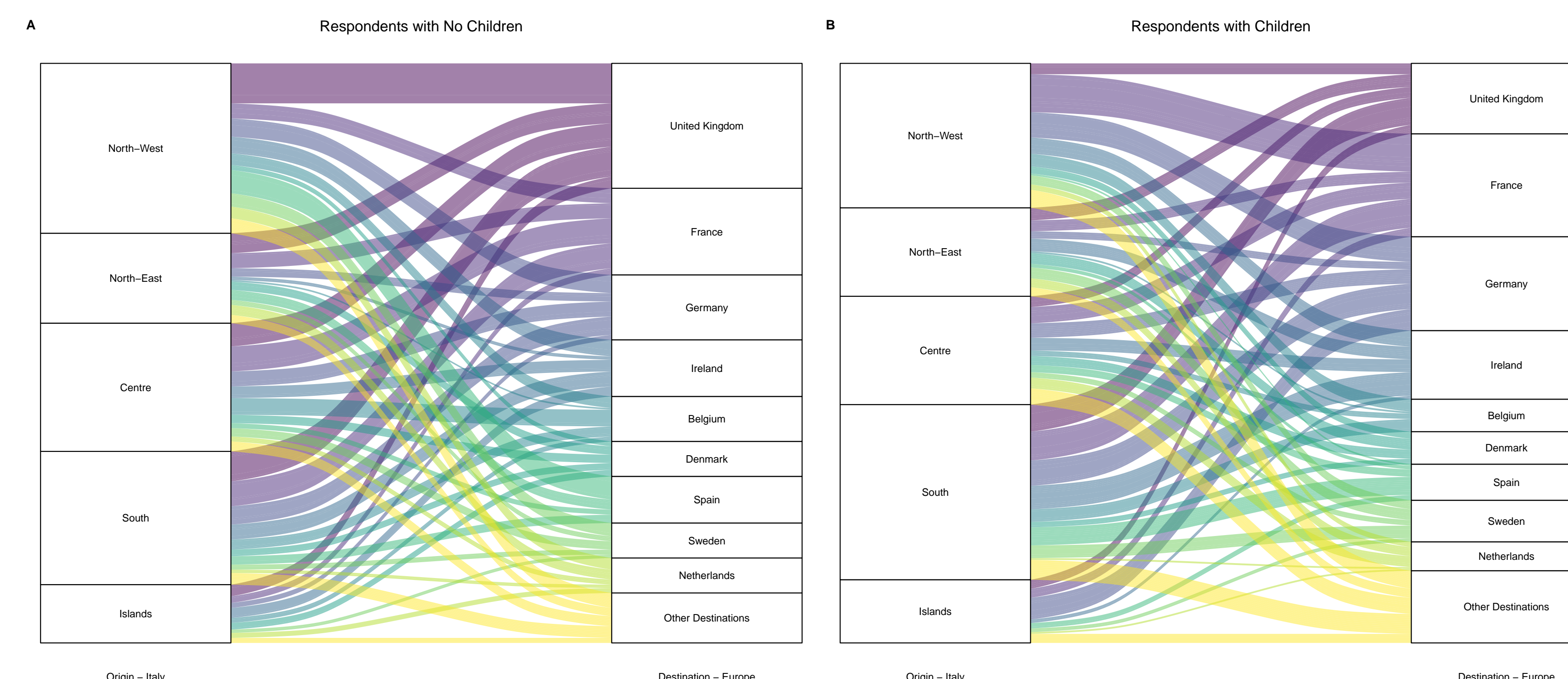


Figure 3: Data from MammaPack Survey respondents illustrating the origin by macro Italian regions and destination at the national level to European countries. Figure 3 (A) focuses on respondents with no children, while Figure 3 (B) highlights respondents with children.

Summary

- The **distribution** 🌍 of MammaPack customers **mirrors** the distribution of Italian residents abroad according to AIRE data (Figure 1), except in Spain. Italians in Spain, likely due to the similar cuisine, purchase less from MammaPack (Coimbra Vieira et al., 2024).
- Training a model** on survey data enabled us to identify **products for babies** 🍼 (Figures 2), helping us **predict** preferences for all MammaPack customers and **identify** those with **children**.
- MammaPack survey respondents revealed fertility patterns from **Italian macro regions to other countries**: it is interesting to note that **southern** MammaPack respondents tend to leave Italy and **have children abroad** 👶.

Conclusions

It is evident that identifying new parents can be achieved by analysing consumer purchases of **baby food products** at small geographic granularities. This **case study** highlights the potential of consumer data analysis for studying specific migrant populations. **Consumer data** 🛒 could also be **relevant** for **investigating** other **demographic behaviours**.

References

- Anelli, M. and Balbo, N. (2021). Fertility Drain or Fertility Gain? Emigration and Fertility During the Great Recession in Italy. *Demography*, 58(2):631–654.
- Coimbra Vieira, C., Lohmann, S., and Zagheni, E. (2024). The Value of Cultural Similarity for Predicting Migration: Evidence from Food and Drink Interests in Digital Trace Data. *Population and Development Review*, 50(1):149–176. [eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/padr.12607](https://onlinelibrary.wiley.com/doi/pdf/10.1111/padr.12607).
- Lichter, D. T., Parisi, D., and Ambinakudige, S. (2020). The Spatial Integration of Immigrants in Europe: A Cross-National Study. *Population Research and Policy Review*, 39(3):465–491.
- Mussino, E. and Cantalini, S. (2024). Multiple Origins and Multiple Destinations: The Fertility of Immigrant Women in Europe. *Journal of International Migration and Integration*.

Acknowledgements

This research is supported by the Saïd Business School Development Fund (MQD02902.MQ04.01) and the Leverhulme Trust (Grant RC-2018-003) for the Leverhulme Centre for Demographic Science.