Quantifying the impacts of COVID-19 on human mobility from an official statistics perspective – the case of Hungary

Zoltán Csányi, Anna Sára Ligeti, János Novák, Ferenc Urbán, László Zoltán Zöldi

The COVID-19 pandemics put official data producers in a new situation. Decision-makers, researchers, representatives of the media and the general audience demand for high-quality statistics on the wide-ranging socio-economic and demographic impacts of the coronavirus, which – regarding depth and timeliness – had no precedents. Official statisticians mostly reached to satisfy such user needs, backed by innovative solutions both at the level of organization and of data engineering. However, despite the obvious consequences of the pandemics on virtually all aspects of human mobility, satisfactory responses of migration statistics are still scarce. In most part, this is due to the mere statistical definition of migration (linked to the concept of usual residence) that disregards shorter term population movements, thus being unable to capture the fast-changing and diverse universe of geographic mobility. In this paper, a wider sense concept of mobility is applied, which we believe is more appropriate for taking stock of the mobile population in the context of COVID-19. In an attempt of quantifying the impacts of the pandemics on outwards mobility of Hungarian citizens and their return, we carried out interrupted time series analyses and fitted an ensemble machine learning model trained by pre-pandemic time series data of monthly migration flows. First we simulated counterfactual monthly flows for 2020 and 2021 in order to shed lights on how migration patterns would have been evolving, if COVID-19 had not disrupted mobility dynamics. Then we compared such model-based predictions with the actual size of monthly migrations in the same years to make conclusions on how the pandemics affected the mobility patterns of the target population. In accordance with our preliminary results, an immediate shock effect of the coronavirus reduced the size of monthly outflows in the second quarter of 2020, however left intra-annual seasonal trends untouched. The lasting negative effect remained significant in the entire period under consideration. In contrast, changes as regards the levels of return mobility cannot be observed, however the actual seasonal patterns differ from the expected ones: many returners apparently brought forward their homecoming due to the first wave of the pandemics. A further important conclusion we might draw from this study is that a large share of the COVID impacts on migration patterns cannot be revealed unless we develop a more dynamic and more flexible understanding of human mobility.

# Introduction

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Morbi at imperdiet tortor, vitae tincidunt felis. Fusce porta mollis suscipit. Aenean sit amet suscipit quam. Nulla a cursus magna. Nulla et venenatis sapien. Nulla pharetra lorem et dui maximus condimentum. Nam non purus vitae tellus fermentum auctor at vel enim. Nulla viverra eleifend sem, ac dapibus orci pretium non. Maecenas eu cursus eros, ut sagittis dui. Cras lobortis lorem vitae purus tristique laoreet. Morbi volutpat, elit in consequat hendrerit, lectus velit tempus ipsum, laoreet tempus dui risus id quam. Phasellus ex purus, suscipit ut quam sit amet, cursus aliquam mi.

# Data

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Morbi at imperdiet tortor, vitae tincidunt felis. Fusce porta mollis suscipit. Aenean sit amet suscipit quam. Nulla a cursus magna. Nulla et venenatis sapien. Nulla pharetra lorem et dui maximus condimentum. Nam non purus vitae tellus fermentum auctor at vel enim. Nulla viverra eleifend sem, ac dapibus orci pretium non. Maecenas eu cursus eros, ut sagittis dui. Cras lobortis lorem vitae purus tristique laoreet. Morbi volutpat, elit in consequat hendrerit, lectus velit tempus ipsum, laoreet tempus dui risus id quam. Phasellus ex purus, suscipit ut quam sit amet, cursus aliquam mi.

# Methodology

To demonstrate the basic idea of our methodological approach, it might be helpful to reformulate our research question using the terminology of the program evaluation literature. Instead of directly investigating the impacts of the pandemic on migration flows, we could ask the following: How many Hungarian citizens would have migrated between 2021 and 2022 if the restrictions due to the pandemic had not taken place? In this way, we can express the effect of the pandemic as the difference between the actual and a counterfactual outcome, i.e. the volume of migration flows during the investigated period in the absence of the pandemic. Formally, we are interested in

where is the actual number of Hungarian migrants, is the corresponding counterfactual , is the beginning of the investigation, and is the elapsed time since the intervention.

In equation 1, the counterfactual outcome is unknown. Therefore, the key issue of our study is its identification. In program evaluation studies, a common solution to the problem is to replace the counterfactual by observations of a control group unaffected by the intervention. However, finding a suitable control group for Hungarian migrants is unfeasible in the context of the pandemic. Instead, we model the time series of Hungarian citizens’ pre-pandemic mobility and forecast future flows from this model to proxy for the counterfactual.

In general, we model the flow of Hungarian migrants at time as a function of its lagged values. If trained by pre-pandemic data free from the influences of the intervention, forecasts of this model yield estimates for the unknown counterfactual mobility pattern in the years of the pandemic:

There are several features of the above model that we have to address in detail. Firstly, in contrast to interventions in program evaluation studies, the start of the pandemic has no well-defined point in time. Even though the first restrictions targeted at the mobility of Hungarian citizens were effective from April 2021, many migrants may have acted accordingly as soon as the restrictions were introduced. As a result, extensively widening the window of our training data may result in biased estimates of the counterfactual incorporating the influence of the pandemic. After experimenting with the shift of this parameter, our first results indicated that choosing December 2020 as the last training observation minimizes the bias arising from this problem, while it still allows our model to learn the latest changepoints in the time series.

# Results

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Morbi at imperdiet tortor, vitae tincidunt felis. Fusce porta mollis suscipit. Aenean sit amet suscipit quam. Nulla a cursus magna. Nulla et venenatis sapien. Nulla pharetra lorem et dui maximus condimentum. Nam non purus vitae tellus fermentum auctor at vel enim. Nulla viverra eleifend sem, ac dapibus orci pretium non. Maecenas eu cursus eros, ut sagittis dui. Cras lobortis lorem vitae purus tristique laoreet. Morbi volutpat, elit in consequat hendrerit, lectus velit tempus ipsum, laoreet tempus dui risus id quam. Phasellus ex purus, suscipit ut quam sit amet, cursus aliquam mi.

# Conclusion

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Morbi at imperdiet tortor, vitae tincidunt felis. Fusce porta mollis suscipit. Aenean sit amet suscipit quam. Nulla a cursus magna. Nulla et venenatis sapien. Nulla pharetra lorem et dui maximus condimentum. Nam non purus vitae tellus fermentum auctor at vel enim. Nulla viverra eleifend sem, ac dapibus orci pretium non. Maecenas eu cursus eros, ut sagittis dui. Cras lobortis lorem vitae purus tristique laoreet. Morbi volutpat, elit in consequat hendrerit, lectus velit tempus ipsum, laoreet tempus dui risus id quam. Phasellus ex purus, suscipit ut quam sit amet, cursus aliquam mi.