A comprehensive analysis of the individual and regional factors explaining the success in obtaining European research and innovation funding

The recent availability of granular datasets has opened up new possibilities to analyse the role of individual heterogeneity on a wide range of economic and social outcomes. Adding to this growing body of literature, our study makes use of a unique micro-level dataset to explore how the probability of signing a contract under the Horizon 2020 European research and innovation programme (henceforth, H2020) is affected by regional factors and individual characteristics of the applicants.

We combine the information contained in the H2020 database with the following: (1) firm-level balance sheet data extracted from the Orbis dataset; and (2) the European Regional Development Fund (ERDF) dataset that includes all projects co-funded by the ERDF over the same period covered by H2020 i.e., 2014-2020.

The main goal of the study is to present empirical evidence on what affects the probability of receiving H2020 funding, considering both the beneficiaries and the characteristics of the territories in which they are located. The empirical model utilised in the analysiscomprises a system of two equations. The first one is related to the decision to apply for the funding, and the second deals with getting access to it. Due to the richness of the data, we are able to control for potential issues like sample selection and endogeneity by relying on large control groups of entities with similar characteristics that either never applied for the H2020 funding, or did but were unsuccessful. Additionally, we use the information on the financing received under other instruments associated with the ERDF to explore whether receiving the ERDF funds might affect the probability of applying and/or participating in H2020 projects.

The combination of the H2020 and Orbis data allows us to study the importance of regional factors for the probability of succeeding in the H2020 programme, such as the number of applicants in the region, the size of R&D spending, and the importance of the Higher Education sector.

Furthermore, as our dataset has thousands of observations and multiple possible explanatory variables, we apply the least absolute shrinkage and selection operator (LASSO) method to select those explanatory variables that significantly increase the accuracy of the model's predictions.

The results of this analysis will allow us to draw relevant policy and practical implications. Highlighting the individual and regional factors associated with H2020 success may help both firms and policy makers to improve their innovation ecosystem. Future research could explore the role of the top R&D investors (Amoroso and Vannuccini, 2019), or the dynamics of the various waves of the European research and innovation programmes by merging the H2020 data with the FP7 ones which refer to the 2007-2013 programming period (Enger and Castellacci, 2016).Finally, we could explore whether setting specific priorities for the current work programme has had any impact on the number of patents and publications in the areas related to these targets (for example, development of next generation batteries, low-carbon solutions, food security, sustainable agriculture, etc.) (Wanzenböck *et al.*, 2020).

References

Amoroso, S. and Vannuccini, S. (2019). Teaming up with Large R&D Investors: Good or Bad for Knowledge Production and Diffusion? JRC Working Papers on Corporate R&D and Innovation No 05/2019, Joint Research Centre.

Enger, S. G., and Castellacci, F. (2016). Who gets Horizon 2020 research grants? Propensity to apply and probability to succeed in a two-step analysis. *Scientometrics* 109.3, 1611-1638.

Wanzenböck, I., Neuländtner, M., and Scherngell, T. (2020). Impacts of EU funded R&D networks on the generation of key enabling technologies: Empirical evidence from a regional perspective. *Papers in Regional Science* 99.1, 3-24.