Dear Editors,

Please find attached our manuscript entitled 'Using the web to predict regional trade flows: material and immaterial regional interdependencies'.

Despite the importance of interregional trade for building effective regional economic policies, we have very little hard data to illustrate such interdependencies. This paper proposes *a novel research framework* to predict interregional trade flows by utilising freely available web data and state-of-the-art machine learning algorithms. Specifically, we employ the JISC UK Web Domain Dataset in order to extract hyperlinks between geolocated commercial websites in the UK. This dataset, which is a subset of the Internet Archive, includes all the archived webpages under the .uk country code Top Level Domain. We are able to geolocate these webpages by searching the archived web text for the inclusion of a UK postcode. In addition, this dataset also contains the hyperlinks included in the archived webpages. Hence, we are able to aggregate these data and create an interregional network of hyperlinks between geolocated commercial webpages. We then use these data to make out-of-sample predictions of interregional trade flows using a rolling-forecasting framework based on Random Forests. We are able to disaggregate our predictions in terms of industrial sectors, but also in terms of space. In total, our models achieve very high predictive capability and provide a proof of concept that the digital traces economic activities, such as interregional trade, leave behind can help us observe such economic phenomena more accurately and, consequently, inform regional policies.

We confirm that this manuscript is not published elsewhere, is not currently under consideration for publication elsewhere and will not be submitted to another journal for consideration until a decision has been rendered by the editors.

Thank you for your consideration of the manuscript. We look forward to hearing from you.

Kind regards,

Dr. Emmanouil Tranos