Review of: 'A Replication of "The Role of Intermediaries in Facilitating Trade" (*Journal of International Economics*, 2011)', by Jianhua Duan, Xuefeng Qian, Kuntal K. Das, Laura Meriluoto and W. Robert Reed (DQDMR) – E-Journal Economics MS 3353

This paper discusses a replication of the empirical analysis reported in Ahn, Khandelwal and Wei (2011) (AKW). The paper also provides robustness checks of the 2011 results by means of exploiting additional data (using primarily World Bank data and Chinese customs data) and sub-sample variation. The theoretical model that motivates the empirical analysis of AKW (and by implication of DQDMR) is discussed qualitatively by means of a diagram (Figure 1). DQDMR have to collect much of the data themselves but obtained programming code and some data from AKW. DQDMR conclude that AKW can be perfectly replicated but that the results are not robust to changes in available data and sample selection.

Besides the paper I have also read the response by AKW to the replication analysis and the detailed comments by another referee. This facilitates the open discussion that *Economics* encourages.

I am impressed with the amount of effort that has gone into data collection, reproduction of the original results and the robustness analysis. Unlike what has been best practice in many other disciplines for a long time, transparency and reproducibility of empirical work in economics remains often rather poor. Hence recent efforts to improve this aspect of economic research are laudable (see also e.g. Christensen & Miguel, 2018).

DQDMR find that the regressions of AKW can be almost perfectly replicated. I agree with them that this is an impressive result because they had to collect much of the data themselves. Once this fact has been established, the reader will be mostly interested in the robustness checks (hence the substance of this paper is contained in section VII and beyond). DQDMR conclude that using more recent data and cutting up of the sample into sub-samples yields some regressions with statistically different results. DQDMR conclude that this weakens the empirical support for AKW's hypotheses. For example, when splitting the sample of firms into three regions (East, Central and West) the results for East in Table 5 are very similar to those of 3(a) and 3(b) in Table 2 but those for Central and West are not. I am not at all surprised by this finding because East accounts for 92% of all observations. The fact that the regression for the 8 percent of the data originating from Central and West yields the opposite conclusion does suggest that there is something going on (AKW note that the fact that intermediaries may source products from outside their region may be one reason for the "anomaly", but there may be additional causes of heterogeneity), but it does not refute the evidence for China as a whole. If the geographic identifier in the customs data plays a role I would have expected to see both theoretical predictions in this paper of why this matters and a more refined spatial analysis (e.g. splitting East further into sub-regions). The purpose of replication should ultimately be to act as a stepping stone for new primary analysis.

I also agree with referee #1 that a full robustness analysis should go beyond cutting up the data or updating data, but include an assessment of the model specification: are there issues of selection effects, endogeneity and omitted variable bias?

For example, it would seem to me that given that firm and/or establishment data are available at different points in time (as DQDMR show), this line of research would really benefit from a panel data approach with firm fixed or random effects rather than OLS estimation.

Minor points

Section II: I am surprised that there is no explicit discussion of the role of intermediaries in information flows and networks, while – for example - the Chinese diaspora play an explicit role in Table 3.

While some readers of a replication study may also read the original paper, it would be really helpful to readers to ensure that the replication study is self-contained. I would therefore suggest that the mathematical formulae and symbols in Figure 1 and in Section III are dropped, because there are not necessary for describing the theory in qualitative terms and the symbols are not explained in the text in any case.

Reference

Christensen, G. & Miguel, E. (2018) Transparency, reproducibility, and the credibility of economics research. *Journal of Economic Literature* 56(3): 920–980.