

Response to Reader 2's Comments on

“Export activity, innovation and institutions in Southern European nascent entrepreneurship”

Thank you very much for your constructive and thoughtful comments, which are much appreciated. I am happy to revise the paper as described in the responses below.

1. *The article's contribution is somewhat weak. There are countless studies investigating the relationship between technological innovations and both export participation and export intensity. Community Innovation Survey is often used (sometimes combined with structural business statistics) but also the World Bank Enterprise Survey, EFIGE and ESEE (Encuesta Sobre Estrategias Empresariales) (look for reference listed below). Authors must quote these studies, particularly those who employ data for Spain. The methods (ordered logit model and logit) are also standard. The data set only allows to estimate correlations based on pooled data as individuals cannot be tracked over time. The main finding that there is a link between export and technological innovation and the training of the founder is well known. Why do we need another study here?*

Response: You are correct that other databases are available. GEM APS and NES indicators have the advantage of picking-up newly formed businesses, which have a different character to older ones. It is also harmonized across countries and it contains policy variables in the same topics as the individual-level variables. As a drawback, it is cross-section, as you highlight. This explanation, together with a recognition of the other studies that you kindly cite, has been added to the text.

2. *To make the paper interesting, I suggest that separate estimates be made for micro (0 -1 employees) and small enterprises (10 to 49 employees). The distinction between manufacturing and services would also be interesting, as export participation is low here, although the tradability of services has increased.*

Response: The individual-level regressions have been repeated for the subsamples you suggest. The estimation results tables have been placed in an Appendix as well as complemented with an explanation in the newly created robustness checks section (section 5).

Size subsamples: Note that 47.25% of the businesses belong to the micro firm subsample, whereas only 6.39% can be found in the small firm subsample. As a consequence, the results for the former subsample replicate by and large those presented in the paper, but for the latter subsample only the number of owners remains relevant due to the small size of the subsample.

Sector subsamples: Note that 87.76% of the businesses belong to the services subsample, whereas only 12.24% can be found in the manufacturing subsample. As a consequence, the results for the former subsample replicate by and large those presented in the paper, but for

the latter subsample only selling a new product and being motivated by just maintaining income remain relevant due to the small size of the subsample.

3. *Conceptual part: The link between exports and technological innovation is not very well motivated. Please derive the research questions or hypotheses. Motivate whether and why the relationships differ in terms of the level of innovation activities (new versus very latest and some versus all). It seems that the definition differs from that of other surveys.*

Response: Please note that Figure 2 presents a model of export activity in entrepreneurship. All variables used are explained in Section 3 and in the Appendix.

4. *The results are based on the ordered logit and binary logit models. The Heckman selection model is mainly mentioned in the footnote. I believe that the Heckman selection model is not suitable. The right variable is the export intensity measured on an ordinal scale. A zero in this case it is a true zero, it is an observed value; this means that the individual (establishment) is not exporting. The Heckman model is suitable when there are not observable values because of truncation, e.g. when the wage is not reported because the person is not working. You need exclusion restrictions to identify the Heckman selection model.*

(...) Heckman selection model not appropriate here. The zero-inflated ordered probit model would be appropriate: Harris, M. N., & Zhao, X. (2007). A zero-inflated ordered probit model, with an application to modelling tobacco consumption. Journal of Econometrics, 141(2), 1073-1099.

Response: Please note that in Section 3 it is said that: “The regression models for export propensity and for export intensity **may be** related by means of a second stage ordered probit for positive export intensity with a Heckman selection correction estimated as a first stage binary probit for export propensity (see De Luca and Perotti, 2011). However, in this sample, **the correlation between the errors of the two equations is not significant, so the selection model is not required.** Therefore, **the export intensity and propensity models are estimated separately**, with the former being estimated as ordered logit, and the latter being estimated as a binary logit.”

The zero-inflated ordered probit is available in STATA16 as *zioprobit*. The individual-level regressions have been repeated for the alternative method you suggest. The estimation results tables have been placed in an Appendix as well as complemented with an explanation in the newly created robustness checks section (section 5). Having a new product or a new technology influences the existence of zeros. Variables such as the education level of the entrepreneur, knowing other entrepreneurs, the number of owners of the business, or being a woman, do not impact on the existence of zeros. Therefore, the zero-inflation is explained by having a new product or a new technology. Note, however, that by placing those variables in the zero-inflating equation, they disappear from the exporting equation to avoid multicollinearity, and this changes the model. Nevertheless, the remaining variables in the exporting equation do not change their behavior. As a conclusion, it can be said that the use of the zero-inflated ordered probit reveals a selection effect for having a new product or a new

technology, such that these variables explain the decision to export. This result does not change the main conclusions of the paper.

5. *I'm not impressed by the data used. Is it the first analysis using GEM data? It seems that the data cannot be linked over time. Individual effects cannot be considered, only pooled regressions can be performed.*

Response: GEM APS and NES indicators have the advantage of picking-up newly formed businesses, which have a different character to older ones. It is also harmonized across countries and it contains policy variables in the same topics as the individual-level variables. As a drawback, it is cross-section, as you highlight. This explanation, together with a recognition of the other studies that you kindly cite, has been added to the text.

Note, however, that the full sample is composed of 46.22% nascent entrepreneurs ("respondents (18-64) involved in nascent business, defined as active, expect to be a full or part owner, and **no salaries or wages paid for over three months**") and 53.78% baby business owner-managers ("respondents (18-64) involved as owner and manager in new firms for which **salaries or wages have been paid between 3 and 42 months**"). The survey question for the former group is "What proportion of your customers **will** normally live outside the country?", whereas for the latter group is "What proportion of your customers normally live outside the country?". Although the variable is not forward as such, in fact it has a prospective nature. The individual-level regressions have been repeated and the estimation results tables have been placed in an Appendix as well as complemented with an explanation in the newly created robustness checks section (section 5). Although the main results for new product and new technology are the same for the two groups, there are differences in several variables. Start-ups (nascent) are negatively affected by the number of competitors, whilst owner-manager young businesses are positively affected. Also, women are negatively affected in start-ups but not in young businesses, whereas the number of owners has a positive effect only in the latter group. Tertiary education increases exports only in the start-up group. In this group, greater independence goes against exports, but just maintaining income fosters them, whilst in the owner-manager group increasing income makes exporting less likely. These differences can be interpreted as prospective versus contemporaneous.

6. *I don't understand why you're using data until 2010. I have checked the GEM website. The data is available until 2015 (GEM 2015 APS Global National Level Data; GEM 2015 NES Global National Level Data). These full datasets are only made available to the public 3 years after data collection. I suggest to update the analysis.*

Response: You are correct that more recent data is available. GEM APS and NES indicators are available until 2018 for Greece, Italy and Spain, and until 2016 for Portugal. The full datasets are only made available to the public 3 years after data collection, so that 2015 is the last year available online. In the data for 2011-2015, the country structure is not very different: Spain 81.48%, Greece 7.69%, Portugal 6.17%, Italy 4.66%.

The individual-level regressions have been repeated for 2011-2015 and the estimation results tables have been placed in an Appendix as well as complemented with an explanation in the newly created robustness checks section (section 5). Note that some adjustments had to be made as some variables that were reported as continuous are now reported in intervals and the dummy for woman entrepreneur had to be removed for the reason explained below. Besides, the export intensity intervals differ and the industry classification is not the same (now Industry ISIC version 4, 1-digit is used). As such, it does not make sense to simply extend the sample and the regressions for 2011-2015 have been added separately.

Regarding the descriptives, for the post-recession period it can be said that 65.42% of the businesses does not export, whilst 5.15% exports more than 75% of sales. There are 57.44% of businesses with 2 owners and 25.14% with 3 owners. Only 4.47% of the surveyed entrepreneurs are women. A very recent (recent) technology is used by 13.10% (20.07%) of the businesses, whilst 36.13% of the businesses sell a new product. The same product is offered by many/few/none of other businesses for 53.57%/35.78%/10.66% of the sample. Regarding the education level of the entrepreneurs, 33.05% / 60.74% / 2.90% has primary-level / secondary-level / tertiary-level studies. To increase income / gain independence / just out of necessity is the motive for setting up a new business for 21.94% / 16.57% / 34.23% of the entrepreneurs surveyed. To have known other entrepreneurs was important in setting up a new business for 30.41% of the sample.

From the preliminary results, it can be seen that the current number of jobs, job growth and the number of owners influence positively the probability of exporting, whereas having a new product or a new technology now contribute negatively towards exporting.

7. The sample is not representative for Italy and Portugal. Re-run the analysis using data for ES and GR only.

Response: The individual-level regressions have been repeated for the subsample you suggest. The estimation results tables have been placed in an Appendix as well as complemented with an explanation in the newly created robustness checks section (section 5). Note that the number of observations for Spain and Greece correspond to 93.87% of the total data used in the paper. Therefore, repeating the estimations for these two countries does not change the results. This also means that there are no outliers in Portugal and Italy that might have been driving the results. In these countries, the relevant covariate is the existence of new technology. However, as you point out, the subsample is small and conclusions should be drawn from the full dataset.

8. I find it difficult to include expected job growth and motives on the right hand side of the equation. Better to include lagged variables.

Response: As said above, the GEM dataset is cross-section. Note, however, that the full sample is composed of 46.22% nascent entrepreneurs ("respondents (18-64) involved in nascent business, defined as active, expect to be a full or part owner, and **no salaries or wages paid for over three months**") and 53.78% baby business owner-managers ("respondents (18-64) involved as owner and manager in new firms for which **salaries or wages have been paid**").

between 3 and 42 months"). The survey question for the former group is "What proportion of your customers **will** normally live outside the country?", whereas for the latter group is "What proportion of your customers normally live outside the country?". Although the variable is not forward as such, in fact it has a prospective nature.