

Review of: "The Relationship Between Overpopulation and Crime Rates in San Andres Island"

Jon Olaskoaga Larrauri¹

¹ University of the Basque Country UPV/EHU

Potential competing interests: No potential competing interests to declare.

It is always a pleasure to read the work of young researchers who are striving to take their first steps in an academic career. This is a stage of learning that greatly benefits from the guidance and support of instructors and colleagues, and one that, over time, is fondly remembered. With this review, I will endeavor to contribute as much as I can to this formative process.

In my opinion, the paper represents an initial approach to an interesting topic, but it has not yet reached the quality standards required by leading journals. Below, I offer some suggestions, most of which correspond to the empirical aspect of social research.

1. It is always recommendable to make an effort to distinguish between the general approach of a doctoral research project and the procedures and findings presented in a specific paper derived from such research. In this article, the research is presented as a *case study*. Case studies are characterized by gathering information from various sources to contextualize the phenomena being studied and employing diverse approaches to access empirical evidence (which usually combine quantitative and qualitative research techniques: interviews with key informants, surveys, life histories, analysis of official documents, statistical analysis of key variables, etc.). The strength of a case study lies in its ability to avoid isolating a single aspect of social reality, instead considering that aspect within the complexity of its social context. Regardless of the approach the author employed in their doctoral research, it is difficult to interpret the content of this paper as constituting a case study. The paper is more accurately described as a quantitative and rather decontextualized analysis of a specific aspect of the social reality of San Andrés Island.

2. In empirical research, it is particularly crucial to identify the relationship between the variables in the theoretical framework and the indicators used for an empirical approach to these variables. In the paper, for example, the theoretical framework addresses the effects of "*overpopulation*"; however, the indicator employed in the model is "*population density*". Some questions that the author should, in my opinion, address include: Is population density the most appropriate indicator for overpopulation? What are the advantages and disadvantages of using this indicator?

3. In quantitative empirical research, it is advisable to report the source of the data used to calculate the indicators, but also how these indicators are calculated, including the scale and units of measurement. For instance, in socioeconomic research, there are several indicators referred to as the "*poverty rate*"; which one is used in the paper?

4. When identifying statistical methods, it is recommendable to use standard expressions that readers familiar with

statistics can correctly interpret. For example, “control variables” is not a method of analysis in itself, whereas “time series analysis” is a term that some readers might interpret—in the context of a paper like this—as referring to the application of ARIMA and ARIMAX models, which, as far as I understand, are not used in this paper.

5. Regardless of the econometric models the author uses to identify relationships between variables, it is always advisable to provide some descriptive statistics informing the reader of the most basic and immediate realities. Throughout this paper, I have not found information on the population density in San Andrés Island, either in isolation or in comparison with other regions in Colombia.

6. When econometric methods are employed, it is necessary to describe the models used as explicitly as possible, including information on the number of observations and the estimation methods. For instance, in this paper, the reader may find it difficult to determine whether a time series has been used—meaning one observation per variable and per year (which would imply only 14 observations)—or whether a panel data set has been used, meaning as many observations per variable per year as there are territories on San Andrés Island. I do not find—though I admit this could be due to a misreading on my part—any information in the paper regarding this matter. Certain parts of the text suggest that panel data might have been used, such as when the geographical breakdown of the island is mentioned in the “Limitations and Mitigating Strategies” section. However, there is no information about the number of territories considered in this hypothetical panel, nor whether fixed effects or, alternatively, random effects have been used in the model estimations.

7. When econometric methods are employed, the presentation of results must be handled with a certain degree of skill, as it is almost always necessary to balance the effort of synthesis with a comprehensive description of the findings. For instance, the models should be presented along with values that indicate their overall significance. When there is a risk that the quality of the estimates may be affected by issues such as autocorrelation or multicollinearity, these risks should be explicitly addressed. If methods are used to test the robustness of the results—such as the instrumental variable mentioned by the author in the “Limitations and Mitigating Strategies” section—the results of the adjustments that include these instrumental variables should also be presented.

8. In the design of tables and graphs, it is essential to add as much information as necessary for readers to interpret each of the elements they contain. This includes titles with information about the indicators represented, units of measurement, etc. I must admit that I have found it difficult to interpret almost any of the graphs presented in the text. For example, Figures 5, 6, and 7 show the cross-values of the crime rate with each of the regressors in the econometric model. First, does the fact that there are only four observations mean that the model's N is indeed 14? Second, the graphs display an almost perfect correlation between the crime rate and each regressor, but this contradicts the data in Table 4, doesn't it? Finally, the relationship between the crime rate and the poverty rate, and between the crime rate and the unemployment rate, is inverse. Isn't this result counterintuitive in the context of the research? Doesn't this require a comment from the author?

9. When the conclusions include an assessment of the implications for policymakers, I would recommend outlining what types of policies are recommended (or most recommended) in light of the research findings, rather than simply urging

policymakers to consider these findings when designing their policies.

10. Finally, please do not include bibliographic references that have not been previously cited in the text, especially if their relevance to the paper is unclear or if they solely serve to cite the author's own work.

I trust that these suggestions may be of some value to the author, both in terms of improving this draft and enhancing their research career.