Gmail Antonio Paez <paezha@gmail.com>

Fwd: Environment and Planning B: Urban Analytics and City Science - Decision on Manuscript ID EPB-2022-0134

Fernando A López Hernández <f8l5h9@gmail.com> Tue, Jun 14, 2022 at 8:18 PM
To: Antonio Paez <paezha@gmail.com>, Pelayo Gonzalez Arbues <pgarbues@idealista.com>, David Rey Blanco <drey@idealista.com>

Hola a todos,
Finalmente tenemos decision del primer paper sometido en EPB. Ver abajo.

He leido por encima los comentarios del referee y no plantea nada insalvable. Al contrario, creo que es relativamente fácil introducir algunos cambios en el paper y elaborar una respuesta.

Se me ocurre que podriamos reunirnos en Teams cuando os venga bien y distribuir trabajo. Yo tengo bastante disponibilidad, así que quedo a vuetra diposición.

Abrazo a todos

Fernando

Catedrático de Universidad
Departamento Métodos Cuantitativos e Informáticos
Universidad Politécnica de Cartagena
Facultad de Ciencias de la Empresa
C/ Real, 3
30.201 Cartagena (Murcia) Spain
Tel.: +34 968 325 619
E-mail: fernando.lopez@upct.es

Webpage: http://metodos.upct.es/falopez

Inicio del mensaje reenviado:

De: Shalini Lakhera <onbehalfof@manuscriptcentral.com>
Asunto: Environment and Planning B: Urban Analytics and City Science - Decision on Manuscript ID EPB-2022-0134
Fecha: 14 de junio de 2022, 10:32:50 CEST
Para: fernando.lopez@upct.es
Reenviado por: <fernando.lopez@upct.es>
Responder a: epb@sagepub.co.uk

14-Jun-2022

Dear Dr. Lopez

Manuscript ID EPB-2022-0134 entitled "Using machine learning to identify spatial market segments. A reproducible study of major Spanish markets" which you submitted to the Environment and Planning B: Urban Analytics and City Science, has been reviewed. The comments of the referee(s) are included at the bottom of this letter.

One reviewer recommended publication outright (after copy-editing), and another referree has been critical of your paper, and have suggested major revisions to your manuscript. Therefore, I invite you to respond to the referee(s)' comments and revise your manuscript. Your revised manuscript will be returned to the

referees for further review. Please note that referees sometimes harden their stance when reviewing the second time around.

To revise your manuscript, log into https://mc04.manuscriptcentral.com/epb and enter your Author Centre, where you will find your manuscript title listed under "Manuscripts with Decisions." Under "Actions," click on "Create a Revision." Your manuscript number has been appended to denote a revision.

You may also click the below link to start the revision process (or continue the process if you have already started your revision) for your manuscript. If you use the below link you will not be required to login to ScholarOne Manuscripts.

*** PLEASE NOTE: This is a two-step process. After clicking on the link, you will be directed to a webpage to confirm. ***

https://mc04.manuscriptcentral.com/epb?URL_MASK=946d9e41074a44518e5c745dea164aa2

You will be unable to make your revisions on the originally submitted version of the manuscript. Instead, revise your manuscript using a word processing program and save it on your computer. Please also highlight the changes to your manuscript within the document by using the track changes mode in MS Word or by using bold or coloured text. Also, please upload a separate document by including your response (detailed) to the reviewers comments. Please note the requirements for figures in the author guidelines at http://www.envplan.com/bauthors.html and upload the correct files.

Once the revised manuscript is prepared, you can upload it and submit it through your Author Centre.

When submitting your revised manuscript, you will be able to respond to the comments made by the referee(s) in the space provided. You can use this space to document any changes you make to the original manuscript. In order to expedite the processing of the revised manuscript, please be as specific as possible in your response to the referee(s).

IMPORTANT: Your original files are available to you when you upload your revised manuscript. Please delete any redundant files before completing the submission.

Because we are trying to facilitate timely publication of manuscripts submitted to the Environment and Planning B: Urban Analytics and City Science, your revised manuscript should be submitted by 12-Sep-2022. If it is not possible for you to submit your revision by this date, we may have to consider your paper as a new submission.

Once again, thank you for submitting your manuscript to the Environment and Planning B: Urban Analytics and City Science and I look forward to receiving your revision.

Yours sincerely

Dr. Levi Wolf
Editor, Environment and Planning B: Urban Analytics and City Science
epb@sagepub.co.uk

Referee(s)' Comments to Author:

Referee: 1

Comments to the Author
Environment and Planning B. Urban Analytics and City Science
Report on
Title: “Using machine learning to identify spatial market segments. A reproducible study of major Spanish markets”

The paper introduces a two-step approach which identifies market segments guided by machine learning algorithms. The first step allows to identify non-orthogonal and non-linear submarket boundaries, i.e. they are not limited to rectangles. The second step includes the spatial hedonic regression.
The paper addresses an interesting topic and provides a substantial contribution by introducing a two-step approach for the pricing of real estate. I have several comments mainly related to the methodology which

might be helpful for the authors in order to make the paper publishable in the journal of Environment and Planning B. Urban Analytics and City Science.

Major comments:

1. The authors claim that machine learning techniques are often black boxes generating results which are often counterintuitive and estimates which are not stable enough to support, e.g., the decision making of commercial banks in their lending process by estimating a correct collateral value. Obviously, their approach avoids this issue by simply using classification tree technique in the first step. However, there are causal machine learning approaches nowadays, which overcome such problems (at least to some extent) (see, e.g., Knaus, Lechner, and Strittmatter (2021)). I believe that the paper would benefit from such a discussion.

2. On page 5, the authors highlight the advantage of their modeling strategy in identifying market segments compared to the one proposed by Füss and Koller (2016). It is not entirely clear to me why the use of prices is a better choice. The use of residuals might be adequate as well if the model provides causal effects (see previous comment). In other words, what kind of information about the location are incorporated into the price which we do not see in the residuals. I believe that this essential part needs a more detailed discussion.

3. The empirical analysis only includes cross-sectional data for 2018Q4. However, it would be interesting to see how homogeneous neighborhoods change over time and space, e.g., due to gentrification. I also do not understand why the authors include the three cities which might by nature be more homogenous than smaller cities or rural areas (or high versus less populated areas). How do they differ in terms of city size, average income, local gdp etc.? It seems that some of the coefficients change sign among the cities. For instance, why does the price decreases when the number of rooms increases in Madrid? Finally, I am wondering how the authors estimated the spatial model with matrices of size 44,270 x 44,270 etc. I assume that the authors use a concentrated log-likelihood approach.

4. I like the horse race among the different models including the spatial model. However, in case of the spatial autoregressive model (SAR), the exogenous variables can no longer be interpreted as causal effects. Moreover, it is not clear to me how this spatial specification affects the estimate parameter for the market segments. In particular, there might be a kind of interaction between the spatial weight matrix and the indicator for the market segments. I would like to see a discussion on how the estimates are influenced in their efficiency and consistency. In addition, the authors should test different spatial weight matrices such as distance- and neighborhood-based matrices. For instance, in case of Barcelona and Madrid the six-nearest neighbors definition might not be appropriate because there are areas with missing neighbors in between city areas, which leads to a meaningless long distance to the next direct neighbor.

Minor comments:

There are several typos in the manuscript. I would suggest sending the paper to a professional proofreader.

References

Michael C Knaus, Michael Lechner, and Anthony Strittmatter (2021): Machine learning estimation of heterogeneous causal effects: Empirical Monte Carlo evidence, The Econometrics Journal 24(1): 134–161.

Environment and Planning B.pdf
110K