

Response to 2021-01-25-jue

Author 1

Author 2

May 27, 2022

1 Summary

Comment: This paper contrasts changes in outcomes between 2001 and 2011 for areas where public housing projects were constructed in Gauteng Province, South Africa, compared to areas where planned projects were not constructed. Gauteng contains Pretoria and Johannesburg, with the projects clustered around the latter. The program under which the projects were constructed began in 1994. There are two main novel findings:

Reply: *Just to note, in Data Section 3, we clarify that “We locate projects using a map of housing projects for the Johannesburg metro area (excluding Pretoria) obtained from the Gauteng City Regional Observatory”. We also include a footnote in the introduction “Our data does not cover Pretoria.”*

1. Along with the expected formal constructions, areas within project boundaries also saw large increases in informal construction, including of “backyard houses” built on formal plots.
2. There is some evidence of small spillover effects of the projects outside their boundaries, which raise the socioeconomic status of the surrounding neighborhoods. On formal house transaction prices, the order of magnitude is similar to those found for low-income areas in Diamond and McQuade (2019) although the results in the present paper are not statistically significant.

2 Long-form comments

Comment: Let me start by saying that the paper makes a distinct contribution to the literature, which will benefit from a greater understanding of the effects of public housing in a developing country setting. To this end, I would like to see more direct comparison with the results on spillover effects of public housing construction in high-income countries, reviewed for example in Diamond and McQuade (2019). Outside of the effects on informal outcomes, how do the results differ and what does this tell us?

Reply: *We appreciate the recommendation to better link to the literature. We updated our discussion of price effects in the intro as follows: “We estimate that prices for formal housing transactions increase by 9% between 0 and 0.5 km from project footprints although this result is statistically indistinguishable from zero. This effect is similar to previous price estimates in the literature, which range from 6.5% to 15% in low-income US neighborhoods to 12% in a middle-income neighborhood in Uruguay. Footnote: {See Diamond and McQuade, 2019 and Baum-Snow and Marion, 2009 for US Low Income Housing Tax Credit estimates. See Gonzalez-Pampillon, 2022 for public housing estimates from Uruguay.} Our effect differs from the 2.5% decrease in prices observed in high income US neighborhoods. Footnote: {Diamond and McQuade, 2019 }. Given that South African housing projects are located in relatively low-income neighborhoods, our price effects are largely consistent with the literature. ”*

Comment: Because the authors provide regression results with the fine building categories as separate outcomes, I would like to see more validation (or at least description) of the data. I agree that aerial photographs are quite informative for whether a given building is constructed with formal or informal building materials, but I am not convinced it is easy to ascertain whether a building contains firms, how many firms there are, and whether the firms are formal as suggested by the first two columns of Table 8. I would like the authors to provide the details of the validation exercise they conducted correlating the GeoTerraImage data on number of informal houses with the figures from the census. I'm not sure whether to consider the reported correlation coefficient of 0.85 large or small.

Reply: *Great idea to better validate the building data. We have success validating the informal/formal housing measures and less success with the business measures.*

- *We are able to reasonably validate building and census measures for formal and informal housing. Validation exercises are described in the Data Section 3 finding strong correlations. We note that our correlation estimate is different from the previous version because we are now calculating correlations at the level of the enumerator area to make the census and building data as comparable as possible.*
- *We also find very similar results using both housing measures given by Appendix Table 26 and discussed in Section 8.*
- *We have less success validating business and utility measures, and include caveats in the result interpretations. Validation exercises (and limitations to the exercises) are described in the Data Section 3.*

Comment: For the formal/informal contrast in Table 3, the authors could show that results are robust to using census measures as outcome variables. Brueckner et al. (2019) suggests that there is some social data on backyard housing, I don't suppose the distinction is present in the census?

Reply: *Thank you for the suggestion. Luckily, the distinction is present, and we find similar results as described below.*

- *The census identifies (1) "Informal dwelling (shack in backyard)," (2) "Informal dwelling (shack not in backyard, e.g. in an informal/squatter settlement or on a farm)," and (3) "House/flat/room in backyard" as given by the new footnote added in Section 3.*
- *We find very similar results using both housing measures given by Appendix Table 26 and discussed in Section 8.*

Comment: For the firm variables, I'd like to see evidence that they track with employment numbers from the census. This more-involved validation would be valuable as more researchers use the GeoTerraImage data with Brueckner et al. (2019), at least, having already done so.

Reply: *Great suggestion, which we tried, but found low correlations possibly because people live and work in different locations. We include a full discussion in Data Section 3 and caveat our results interpretations accordingly.*

Comment: I'd also like some clarification on what is/isn't known about each housing project. Analysis is at the level of cells of a 100 m 100 m grid imposed on the area of interest, which the authors refer to as plots. The effect of project construction on number of formal houses in a given project area between 2001 and 2011 is presumably known to project administrators, as is the effect on construction of public service buildings. If the authors accessed this information they could, for instance, estimate the average across projects of the number of formal houses built per plot on plots contained within project areas, where $()$ indicates the cardinality of the argument and I've abstracted away from the fact that some plots are not fully inside or outside a project area. As I understand it, this is effectively what $()$ should measure in equation 1 when the outcome variable is number of formal houses. For the net effect of being in a constructed project, the authors could estimate $()$ for unconstructed projects. The only statistical uncertainty comes from estimation of $()$

so this result would be more precise than what the authors report. Can the authors follow this approach? If not, is it because program administrators do not make the information available (why not? Could it be that fewer houses were constructed than should have been?), or because they don't have records?

Reply: *Thank you for the detailed empirical suggestions, but unfortunately, we face many data limitations described below.*

- *In the Data Section 3, we clarify that: “With this project definition, we do not know the dates when projects were started or completed. While the majority of projects were likely to have started construction after 2001 consistent with a broad expansion in housing programs in the early 2000’s, the data may also include some projects that were already completed by 2001.”*
- *Unfortunately, reliable data on the number of units constructed and dates of construction are not available. We have reached out to administrators to no avail, and we have tried linking project names to government treasury records, but different naming conventions mean that this exercise did not yield any reasonable results. One explanation is that government housing is an intensely political issue in South Africa.*
- *We reframed our estimates of the effect of project construction on formal housing in the project footprint to be in terms of measuring the projects themselves (instead of estimating any exogenous effects). See Measuring Housing Project Construction Section 6.1 for a detailed discussion.*

Comment: In Figure 1, I notice that many unconstructed project areas border constructed project areas so that according to the authors’ own analysis the Stable Unit Treatment Value assumption will not hold for the within plot footprint analysis because the unconstructed plots will be subject to spillover effects. As a robustness check, I would like to see the analysis of at least Table 1 repeated when excluding plots where non-constructed and constructed areas are close, recognizing that these areas are somewhat geographically distinct. Since spillovers appear to be small, I do not expect large differences in results.

Reply: *Excellent suggestion for a robustness test. We find that the main results remain largely unchanged excluding any projects that are within 2 km of any other projects. See new Table 23 and Robustness Section 8*

Comment: I would also like to see the welfare analysis in section 9 removed altogether. The authors cannot treat the actual construction of the public housing itself as the action of a profit-maximizing developer, which is implied by including direct effects on formal housing in equation 6. The authors imply that there is a first-stage relationship between construction costs and steepness but as far as I can tell, they do not have variation in construction cost so it isn't clear how γ in equation 6 is being estimated.

Reply: *Helpful suggestion, and we have removed it.*

3 Short comments

Comment: The authors claim that their measure of plot exposure to the projects is novel since it measures exposure by the fraction of the project area that intersects the plot itself and areas around the plot centroid instead of distance to the nearest project as in Diamond and McQuade (2019). I do not think the measure is novel. At the very least, Autor et al. (2014) have a similarly-continuous measure of exposure to rent controlled buildings in Cambridge, MA, although it is based on number of units rather than area. Gechter and Tsivanidis (2020) contains a similar measure with that paper being developed over the same time period as this one. Maybe the authors mean the measure is novel within the literature on spillover and direct effects of public housing development?

Reply: *Thank you for the helpful cites. We removed claims that it is novel and added citations to Autor et al. (2014) and Gechter and Tsivanidis (2020) in the measurement section 5.1*

Comment: I was a bit confused about the area covered by the public housing project data. Am I right that it is all of Gauteng province and therefore, there was no public housing constructed under the program in Pretoria?

Reply: *Thank you for identifying the lack of context. In the Introduction, we added a footnote clarifying “Our data does not cover Pretoria.” In the Data Section 3, we also clarify that the data does not include Pretoria.*

Comment: I didn’t understand the following interpretation of the coefficients in the “Plot foot-print row of columns 4 and 5 in Table 1: Among informal houses, backyard houses replace non-backyard houses consistent with qualitative evidence that housing authorities cleared preexisting slums before constructing new projects”. Both coefficients are positive and significant whereas the statement made me expect a negative coefficient on informal houses.

Reply: *We have revised our description to the following in Section 6.2 “Columns (2) and (3) disaggregate the total effect into effects on informal backyard houses and informal non-backyard houses. Backyard housing is a common phenomenon in South Africa where households erect informal dwellings within the plots of formal dwellings. We observe a large increase in backyard housing at the same magnitude as the increase in formal housing in Table 3. This finding suggests that almost every formal project house is accompanied by an informal, backyard house constructed on the same parcel of land. We also estimate a relative decline in informal, non-backyard houses in constructed project areas relative to unconstructed project areas. The number of people living in informal housing declines as a result of project construction. These findings are consistent with larger non-backyard houses being replaced by smaller backyard houses.”*

Comment: On page 27, the first instance of the word “which” should be deleted.

Reply: *Great catch, and we have removed the welfare section that contained it.*

Comment: There is a good reason why Diamond and McQuade (2019) and Rossi-Hansberg et al. (2010) do not consider effects on public housing project plots: they are not considering informal activity. I wouldn’t imagine informal activity is first-order in their settings. Perhaps it was in Hornbeck and Keniston (2017)’s historical setting. As such, I would like the authors to moderate their statements in the conclusion about a general need to consider within-area effects of public housing projects.

Reply: *Great observation. In conclusion section 9, we clarify “researchers and policymakers have largely focused on changes outside of policy footprints because policies in developed settings leave little scope for informal activity within their footprints” and emphasize that our results are relevant for developing settings.*