Referee's Report: Assessing effects of nature-based and other municipal adaptation measures on insured heavy rain damages

The study examines the effect of pluvial flood adaptation measures by analyzing insurance claims in two adjacent neighborhoods in the Netherlands before and after the interventions. The authors find that the adaptation measures in their totality cause a decline of 3700 euros per rainy day.

I agree with the authors that we still know too little about the effectiveness of mitigation measures, especially from an economic perspective. Studying the governmental mitigation projects adds to both the academic literature and has high social relevance. The use of insurance data provides a unique and objective measure to study the effectiveness of such measures.

While I think the set-up is interesting, multiple questions regarding the method and economic significance of the findings remain. Addressing these points should largely contribute to the paper, in my opinion. I recommend a revise and resubmit with major revisions.

Methods

I don't think that the visual interpretation around Figure 1 provides convincing evidence that the vital common pre-trend assumption is met. The study would largely benefit from a more formal test of the common pre-trend assumption. For example, through an event study.

Related to this, I wonder if it is likely that the amount of rainfall varies between your treatment and control group? If purely coincidental, the treatment group was affected by a large shower while the control group was not this biases results.

Both your treatment and control groups are affected by the general information campaign. If the effect of this campaign is homogenous for both areas, this should not affect the estimates that you get for the specific projects. However, if this is not the case, your results might be biased. In this regard, I think it is interesting to think about why the government chose to implement mitigation measures in the *Scheldebuurt* and not (yet) in the *Rijnbuurt*? This is ideally random. However, I can imagine that in this case, the firmer was chosen because the initial risk was higher, and thus the information campaign might also be more effective here.

Account for the insurance provider composition. The mitigation measures might be ex-ante observed by insurers who update their fees. Hence, the insurer composition before and after treatment might be altered. If certain insurers receive fewer claims in general (e.g., due to less coverage and user-friendliness of the claim portal), this could drive results.

Can spill-over effects bias the results? Many of the nature-based adaptation measures could also reduce risk in the control areas. For example, if the sewers are linked, increasing water storage capacity in the treatment area also reduces the chance that the water-bearing capacity in the control area is reached.

If I understand correctly, "post" is based on the date when the mitigation program started. However, realizing these projects does not happen instantly, and probably not all projects were implemented at the same time. Leaving out the intervention period should provide a more reliable estimate of the actual benefit, as the current measure likely underestimates the effect.

Time-specific neighborhood-level shocks are controlled for through fixed effects for each month. This addresses general temporal variation but not temporal variation at the neighborhood level. This would require 'month by neighborhood' FE.

Does the control percentage of real estate built before 1945 add new info? Unless a lot of new buildings were added during the study period, all this information should already be absorbed by the PC4 fixed effects. Also, it is unclear how the "address density" and the "value of property" controls are measured. Are those PC4 averages? If not, how can you use singular-level controls in combination with an aggregated dependent variable?

Economic significance

The authors state that their outcomes can be used to make a cost-benefit analysis. I would encourage them to do this already. If costs are difficult to obtain, at least provide some intuition on the benefits. What is the economic significance of saving 3696 euros per rainy day? Is that a lot or a little? How many rainy days are there?

Rather than controlling for rainfall, can you not incorporate this in your measure? E.g., a triple diff with post x treatment x sum of rain? Likely, the benefits from mitigation are not equal for different levels of rain. E.g., the measures only work to a certain limit, or they become even more effective beyond a certain amount of rain. It would be interesting to disentangle this.

The authors claim that studying multiple interventions in a single study is a strength. I would say that knowing about individual intervention contributions adds more value, as it allows for more optimal adaptation given budget constraints. Within the study set-up, it might be difficult to disentangle the benefits of individual adaptation measures. However, I would like to see some discussion on this. Do all measures equally contribute, are some more important than others, or is it exactly the joint effect that makes these measures so effective? Disentangling these added values would help with more efficient adaptation.

It is unclear whether the study utilizes all insurance claims or only those accepted. The latter is probably a cleaner measure. However, it can mask real damages and thus underestimate the net effect or bias results entirely if certain insurers reject more claims than others, and the composition of insurers changes during the study period. Related to this, I would like some reflection on how well insurance claims as a whole measure the effectiveness of mitigation measures. How large is the share of people not making claims despite damages, and what does this imply for your results?