Dear editor,

Thank you for giving us the opportunity for revising our manuscript. Most comments are fair and doable and will definitely make the manuscript stronger. I would however like to hear the editor's opinion on one issue.

Reviewer 2 recommended to reject the paper and you are planning to send the revised paper to reviewer 2 for the second round. Reviewer's 2 main issue with the paper is "how they analyze the data". I wonder if I can and should please reviewer 2. **It would be nice to know from the editor if our revised manuscript has a chance to be accepted for publication in PlosOne, even if Reviewer 2 advises to reject our revised manuscript.**

Reviewer 2 recommends us to "estimate a standard differences-in-differences model using the crisis ASC as the treatment indicator (equal to one if a crisis ASC was set up in the neighbourhood between wave 1 and 2) and the wave as the post-treatment indicator".

This is, however, not possible, because our dependent variable is a binary outcome and our treatment indicator is a continuous variable.

The traditional DiD model (with additional regressors ci) is:

*yit= β1Timet + β2Treati + δ(Timet⋅Treati) + ci + ϵit*  (1)

with *δ* being the DiD estimator and Treati the dichotomous treatment variable.

Formula (1) is equivalent to:

*ΔYi=β1+δ Treati + ϵi* (2)

Since our outcome is a binary variable there is no standard DiD model!! We therefore estimated:

*logit(Pr(ΔYi =1|Treati))= β1 + δ Treati*  (3)

with *ΔYi =1*if the dependent outcome was 1 post-treatment (i.e. wave 2) and 0 if the dependent variable was 1 pre-treatment (wave 1). Our *Treati* variable is the change in exposure the asylum seekers. Formula (3) is the fixed effects logistic regression model for two waves (or, more precisely, the first difference model). Formula (3) demonstrates that the DiD model is a type of fixed effects model because the time constant covariates drop out of the model. With our FE-model, we are thus able to control for time-stable unobserved heterogeneity. This, together with the fact to our treatment is to a large extent random (which we show in our manuscript), we can make strong (but not definite!) claims on causality. Naturally, we are aware that because we have a binary outcome and we use a nonlinear link function we violate the common trend assumption necessary to interpret delta as the DiD estimator. Moreover, we like to point out that our treatment variable *Treati* is not a dichotomous variable, and this also makes why we cannot interpret our effect as the traditional DiD estimator. But we do not claim to estimate a DiD estimator!

As an additional robustness check, I could estimate formula (1) directly for a binary outcome variable. Thus a linear probability model (LPM), while controlling for heteroscedasticity in the error term. And I could dichotomize our treatment effect. The estimate could now be interpret as a DiD estimator but the model is of course less ideal than our original model.

In our manuscript we do not claim to estimate a causal effect. We claim we have a strong design and come much closer to interpreting our finding as a causal effect as previous work. We acknowledge in our manuscript that ‘treated’ respondents and neigbhourhoods may not be completely identical to ‘untreated’ respondents. We address this point in the robustness paragraph. We will, as requested by reviewer 2 make this point more clear. But, we will never be able to prove that there is no unobserved time-varying heterogeneity, thus we will never be able to test the underlying common trend assumption of the DiD model. But once again, we do not claim, nor aim to estimate a DiD estimate.

Thus our response to Reviewer 2 – with the above explanation - would be that:

* to estimate a standard DiD model is not possible.
* That a LPM comes closest to a traditional DiD model but that this is not our preferred model (but we will estimate this model as a robustness check).
* That we cannot prove the underlying common trend assumption of the DiD model. But show that our ‘treated’ and ‘untreated’ respondents are very much alike and can be made more alike by a matching procedure.
* That notwithstanding the above, we can make strong, but not definite, claims on causality.

Given my reading of Reviewer 2 comments I doubt Reviewer 2 will be satisfied with the above. **It would be nice to know from the editor if our revised manuscript has a chance to be accepted for publication in PlosOne, even if Reviewer 2 advises to reject our revised manuscript.**

Thank you very much for your attention and I am looking forward to your reply and, hopefully, to start working on a revised manuscript.

Best regards,

Jochem

time an

ΔYi=β0+β1Xi+ui

Did

Y= β0 + β1\*[Time] + β2\*[Intervention] + β3\*[Time\*Intervention] + β4\*[Covariates]+ε

FD

Y1= β0 + β1\*[Time=1] + β2\*[Intervention] + β3\*[Time=1\*Intervention] + β4\*[Covariates]+ε1

Y2= β0 + β1\*[Time=2] + β2\*[Intervention] + β3\*[Time=2\*Intervention] + β4\*[Covariates]+ε2

Y2 – Y1 = β1\*[Time=2 – Time=1] + β3\*[(Time=2 – Time=1)\*Intervention] + ε2- ε1

<https://stats.stackexchange.com/questions/197749/difference-in-difference-vs-fixed-effect-models>

<https://www.econometrics-with-r.org/13-4-qe.html>

<https://stats.stackexchange.com/questions/89513/difference-in-differences-estimator-for-logistic-regressions>