PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (**http://bmjopen.bmj.com/site/about/resources/checklist.pdf**) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Examining the policy effects of Arizona's 2016 preemption law on firearm suicide rates in the greater Tucson area: An observational study
AUTHORS	Goldstein, Evan; Prater, Laura

VERSION 1 – REVIEW

REVIEWER	Ku, Benson Emory University School of Medicine
REVIEW RETURNED	28-Oct-2021

GENERAL COMMENTS	Authors have written a well written research report that assessed the policy impact of Arizona's 2016 preemption law on firearm suicide rates in the greater Tucson area. They found an increase in Pima County's firearm suicide rate relative to comparison group counties over the same period, while the preemption law did not appear to be significantly related to non-firearm suicide rates in Pima County. Although these results show interesting correlations, I have some major concerns below: 1. Covariates Authors have included several important covariates in their models including age, sex, race, socioeconomic status. However, I would encourage authors to include mental health shortage areas as a covariate since this variable has been shown to be associated with the increase in suicide rates in USA from 2010 to 2018. Controlling for potential covariates would be important for authors to conclude that SB 1487 led to higher firearm suicide rates in Pima County. Ku BS, Li J, Cathy Lally, Compton MT, Druss BG. Associations between mental health shortage areas and county-level suicide rates among adults aged 25 and older in the USA, 2010 to 2018. Gen Hosp Psychiatry. 2021 May-Jun;70:44-50. doi: 10.1016/j.genhosppsych.2021.02.001. Epub 2021 Feb 8. PMID: 33714795; PMCID: PMC8127358. 2. Statistical analysis Would it be possible for authors to use spatial autocorrelation to rigorously control for clustering and dispersion of firearm and non-firearm suicide rates per county in AZ if they have not already done so?

REVIEWER	Tu, Wei Georgia Southern University College of Science and Mathematics, Geology and Geography
REVIEW RETURNED	10-Nov-2021

GENERAL COMMENTS	This study conducted a difference-in-differences analysis to estimate the effect of Arizona's SB 1487 in county-level firearm suicide rates in Pima County using a time-series cross-sectional data from 2014 to 2019. The authors concluded that 1.126 per 100,000 persons increase in Pima County's firearm suicide rate was attributable to the enactment of Arizona's 2016 preemption law, relative to comparison group counties over the study period. This is an interesting study quantifying the impacts of state and local policy on firearm suicide rates in the United States. I have a few questions that would appreciate authors responses 1. What difference will it make if a longer time series data were used to run the model? I understand that only three years of the data on and after the enactment of the law were available when the research has been conducted, but a long time series data before 2016 should be attainable. Figure A1 is good but can more years of data before 2016 be graphed so that at least the trend before the enactment of the law can be more clearly visualized? 2. How long has Tucson's ordinance has been around? Is it possible to compare the rates before and after the enactment of the ordinance? If the ordinance does significantly impact the firearm supply on the market, there should be a rate difference before and after the existence of the ordinance as well. 3. Several well-established covariates in the literature were not included in the models such as veteran status, psychopathologic factors, and behavioral factors. There was also a significant difference in R2 between Model 2 and 4. I am not suggesting that the authors should refit the models with more covariates, but some discussions would strengthen the paper. 4. Will it be helpful to graph the unadjusted firearm and non-firearm suicide rates in Pima county and the group counties? 5. It may also be helpful to add some brief discussion on the difference-in-differences method.

REVIEWER	Riddell, Corinne UC Berkeley, School of Public Health
REVIEW RETURNED	11-Feb-2022

GENERAL COMMENTS	This paper considers a policy enacted by the AZ government in 2016 to prevent the Tucson AZ government from destroying guns they confiscated or were forfeited by their residents. The policy forces the government to resell the guns at a public auction, and in one month >600 guns were sold because of the policy change. Given that the policy only affected Tucson/Pima county (because they were the ones destroying the guns, not surrounding areas in AZ), the authors' propose a difference in differences (DID) design to estimate whether the policy change increased firearm suicides over and above any temporal increases experienced by other AZ counties. The paper is very well-written and engaging. The study is well presented and motivated. I have a few questions about the DID model:

	1) Can you clarify whether you are using linear regression on suicide rate (#suicides/population size) for each county-year as your outcome of interest? If so, have you considered using Poisson or negative binomial regression instead to account for the varying population sizes across counties? I know this means that the interpretation of the interaction term is then multiplicative but was curious about how/why you made this decision. 2) Have you considered including yearly fixed effects in lieu of the post-law enactment period variables, and including a categorical "time_since_policy_change" variable in place of interaction term (so that this variable equals 0 in all untreated states and in pre policy time in AZ, 1 in the first year after treatment in AZ, then 2 in the second year and so on)? In particular, it might be interesting to see if they effect over time was dynamic. It sort of looks like the effect increases with time according to your figure, and would be nice to put some estimates on that. 3) When you describe the covariates you adjust for, it would be helpful to include a sentence about what sorts of variables are confounders in DID analysis (variables that are state and time varying) 4) Either in the methods or the discussion, would also be nice to see a sentence addressing the third assumption of DID — that there was no other change at the time of the policy change that affected Tucson/Pima county and not the other counties. Need to at least mention it as an assumption and perhaps discuss in discussion if there may be something else happening that you know about over that time period. Other comments 5) Round effect estimate and CI to two digits ini abstract and results and tables 6) Describe the effect estimates as risk differences (increase in additive risk of outcome due to policy change) 7) Replace p-values with 95% CIs in results 8) Though interesting, the info on R-squared about the adjusted models is superfluous - don't need a higher R^2 for causal models, just need a model that adjusts/eliminates confounding through the model specification so I suggest removing this from the results 9) I would also suggest removing the sentence in the results about the association between unemployment and firearm availability and suicide rates since it isn't a result related to your study objectives. 10) Rather than talking about 4 models - suggest reframing as 2 models - unadjusted and adjusted - for your primary outcome, and then say you ran the same set of models as part of a sensitivity analysis or negative control test.

VERSION 1 – AUTHOR RESPONSE

Comments from *Reviewer #1*

Comment #1: Authors have written a well written research report that assessed the policy impact of Arizona's 2016 preemption law on firearm suicide rates in the greater Tucson area.

Response #1: Thank you for your thoughtful review. Your comments are much appreciated!

Comment #2: Authors have included several important covariates in their models including age, sex, race, socioeconomic status. However, I would encourage authors to include mental health shortage areas as a covariate since this variable has been shown to be associated with the increase in suicide rates in USA from 2010 to 2018. Controlling for potential covariates would be important for authors to conclude that SB 1487 led to higher firearm suicide rates in Pima County.

Ku BS, Li J, Cathy Lally, Compton MT, Druss BG. Associations between mental health shortage areas and county-level suicide rates among adults aged 25 and older in the USA, 2010 to 2018. Gen Hosp Psychiatry. 2021 May-Jun;70:44-50.

Response #2: That is an excellent suggestion and an important reference. Mental health professional shortage area (HPSA) status is an important "good" covariate missing from our initial analyses (i.e., a covariate not itself influenced by our policy variable).

Although it can take time to request and receive older versions of the Area Health Resource File (AHRF) data from the contractor overseeing the older data, the lead author possessed enough older AHRF data needed to cover observations from 2014-2019. This new data source is now mentioned on p. 7. Incorporating the AHRF data, we made the following changes:

First, we updated our analyses to include a county-level measure of federally-designated mental HPSA status. Of note, all of the county observations included in our study were either partial or full mental HPSAs throughout the study. There were 0 non-mental-HPSAs. There is a long literature incorporating county-level HPSA measures and different approaches to measuring partial HPSAs (e.g., Yu et al.'s recent [2022] paper in the American Journal of Preventive Medicine). If we only consider full mental HPSAs, the variable will be subtracted out of our multivariable analyses as a time- and group-invariant measure. Because this is an important covariate, our mental HPSA measure had two

categories – partial and full – to account for differences in provider supply at the county level.

Tables 1 and 2 have been fully revised. We also updated our description of the results in the manuscript on pp. 11-13. Notably, including the mental HPSA covariate improved the R^2 for Models 2 and 4, helping soak up some of the residual variance in our outcomes.

Second, we cited the recommended Ku et al. paper (2021) and described the mental HPSA variable in the *Covariates* subsection on p. 9.

Comment #3: Would it be possible for authors to use spatial autocorrelation to rigorously control for clustering and dispersion of firearm and non-firearm suicide rates per county in AZ if they have not already done so?

Response #3: Our estimators used clustered standard errors (SEs) at the county level to account for heteroscedasticity and serial (auto) correlation within our geographical units over time. We do this because policy treatment is assigned at the county level in our study design and a source of variation ostensibly comes from differences occurring between Pima and the comparison counties.

If we were to follow a different treatment for our standard errors, such as Beck and Katz's (1995) panel corrected SEs, which can perform well in small panels like ours, we get comparable results to what we report in the manuscript (although the SEs seem less conservative):

Table A. DID estimates with panel corrected SEs from the adjusted model

	1 *Adjusted Model*
Policy variables	
SB 1487 exposure	
Comparison group	Ref
Policy group (Enactment of state law, SB 1487, preempting gun disposal ordinance in Tucson, Pima County)	0.518**
	(0.0007)
Policy enactment timing	
Pre-law enactment	Ref
Post-law enactment	0.297**
	(0.0003)
Policy group x Post-law enactment (difference-in-differences estimate)	1.127**

(0.0009)

Notes: Excerpted policy variable estimates from the fully-adjusted model. ** indicates P < 0.001

However, we follow the treatment assignment rule, embedded in our research design and data collection method, which is recommended for a DID design examining the effects of policy intervention (i.e., Wooldridge, 2002; Arellano, 1987).

References:

Wooldridge JM: Econometric Analysis of Cross Section and Panel Data; 2002; 58:752

Arellano M: Computing Robust Standard Errors for Within-groups Estimators. Oxf Bull Econ Stat; 1987; 49:431–434.

Comments from *Reviewer #2*

Comment #1: What difference will it make if a longer time series data were used to run the model? I understand that only three years of the data on and after the enactment of the law were available when the research has been conducted, but a long time series data before 2016 should be attainable. Figure A1 is good but can more years of data before 2016 be graphed so that at least the trend before the enactment of the law can be more clearly visualized?

Response #1: Adding more years to Figure A1 is an excellent recommendation. We extended Figure A1 to describe the firearm suicide rates from 2005-2019. We also moved Figure A1 to the main manuscript file. It is now labeled Figure 2. This allows the reader to examine the pre-policy trends over a longer period. Importantly for our estimators, the trends are "good" in that they mirror each other but do not converge or cross in the pre-treatment period. The revised figure also underscores how the firearm suicide rates really do not begin to converge until the post-policy enactment period from 2016-2019, when the firearm suicide rates plateau in the comparison group counties but rise noticeably in Pima County.

Unfortunately, due to data limitations, we cannot add these additional years to our multivariable statistical models. The ATF data used to adjust for the important firearm availability proxy is not available before 2014 (Haviland et al., 2021). That said, as a robustness check, by estimating our unadjusted model (Model 1) using the 2005-2019 data, we found similar results to what we report in the manuscript. The β3 difference-in-differences estimate for this test is 1.57 (P = 0.02). Empirically speaking, though, we just need at least two serial observations (2014, 2015) for the policy and comparison groups before the start of the policy to check for the similar preintervention trends assumption (Gertler et al., 2016; Wooldridge et al., 2002).

That said, since we cannot extend the pre-treatment period in our multivariable analysis, we decided to conduct a "placebo test" of the expected policy effects and pre-policy common trends assumption (Gertler et al., 2016). For this placebo test, we performed an additional DID estimation using a "fake" policy group for our outcome of interest. Specifically, we replicated our estimation of Model 2 using Maricopa County for our policy group observations and all other non-Pima counties for the comparison group. Because the 2016 preemption law should not have affected firearm suicide rates in Maricopa County relative to the other comparison counties, the DID estimate (β3) from the placebo test model should not statistically differ from zero. The results of this placebo test are shown in Supplemental Table A1 and described in the manuscript on p. 12. The DID estimate from the placebo test model did not statistically differ from zero at the 0.05 level (β3 = -0.864; P = 0.216; Table

A1). In other words, the 2016 preemption law did not significantly impact firearm suicide rates in the "fake" policy group (Maricopa County), compared to the remaining comparison group counties. If the DID estimate from the placebo test model significantly

differed from zero, the impact would have likely come from some underlying difference in the trends between the two groups. In turn, this would have cast doubt on the assumption of similar pre-policy trends between our main policy and comparison groups.

References:

Haviland MJ, Gause E, Rivara FP, et al.: Assessment of county-level proxy variables for household firearm ownership. Prev Med (Baltim) 2021; 148

Gertler PJ, Martinez S, Premand P, et al.: Impact Evaluation in Practice, Second Edition.

World Bank; 2016.

Wooldridge JM: Econometric Analysis of Cross Section and Panel Data; 2002; 58:752

Comment #2: How long has Tucson's ordinance has been around? Is it possible to compare the rates before and after the enactment of the ordinance? If the ordinance does significantly impact the firearm supply on the market, there should be a rate difference before and after the existence of the ordinance as well.

Response #2: Tucson's ordinance was adopted in 2005 and would have been implemented (training for police, enforced, and marketed) in subsequent years. Evaluating the impact of the original ordinance is beyond the scope of our study; data limitations for key variables (noted above) and other confounding factors extending over such a long period of time preclude us from adequately evaluating the effects of the original ordinance. However, using only the CDC WONDER data, we do see that

the unadjusted firearm suicide rate decreased in Pima County from 10.7 per 100k in 2006 to 8.9 per 100k in 2007 (and 8.8 per 100k in 2008). In contrast, the unadjusted firearm suicide rate did not decrease in the comparison counties from 2006 (12.0 per 100k) to 2007 (11.4 per 100k) or 2008 (12.0 per 100k). We cannot make conclusions about the original ordinance in this study. But, following your previous recommendation, we believe our new Figure 2 helps illustrate the important firearm suicide trends over a longer period of time. We now also note the timing of the Tucson ordinance for the reader in the Figure 2 notes section.

Comment #3: Several well-established covariates in the literature were not included in the models such as veteran status, psychopathologic factors, and behavioral factors. There was also a significant difference in R2 between Model 2 and 4. I am not suggesting that the authors should refit the models with more covariates, but some discussions would strengthen the paper.

Response #3: Thank you for this important comment. We revised our *Limitations* subsection on p. 16 to further underscore how unobserved characteristics not accounted for in Models 2 and 4 may have biased our estimates, imposing limits to causal interpretations of our findings. In additional to firearm availability, we specifically mention county-level veteran population size and unmet need for mental health care, the latter of which we cannot adequately measure. That said, Reviewer #1 made an excellent recommendation to attempt to adjust for county-level mental health professional shortage area (HPSA) to account for differences between the counties in available mental health clinician supply. Including the mental HPSA covariate improved the R^2 for Models 2 and 4, helping soak up some of the residual variance in our outcomes. Finally, regarding the difference in R^2 between Models 2 and 4, the higher R^2 in Model 2 is likely attributable to the amount of variance in the firearm suicide outcome explained by our firearm availability proxy (i.e., the per capita rate of federal firearm licenses). Unlike non-firearm suicide, firearm availability is strongly and consistently associated with firearm suicide in the scholarly literature. Documented elsewhere, firearm availability is very difficult to measure, though the recent study by Haviland et al. (2021) suggests the proxy we used may be the most suitable proxy for county-level analyses.

References:

Haviland MJ, Gause E, Rivara FP, et al.: Assessment of county-level proxy variables for household firearm ownership. Prev Med; 2021; 148

Comment #4: Will it be helpful to graph the unadjusted firearm and non-firearm suicide rates in Pima county and the group counties?

Response #4: Yes, it would be helpful. We made three improvements: First, Figure 2 now shows the unadjusted firearm suicide rates (our primary dependent variable affected by the policy intervention) from 2005-2019. We also moved this figure from the Supplemental Material to the main manuscript file. Second, at the editor's request, we added a new map to illustrate the unadjusted firearm suicide rate for each county

in 2019, helping the reader visualize variation in these rates geographically (Figure 1). Third, although the non-firearm suicide rate variable is only used for our robustness test, we added a new supplemental graph of the unadjusted non-firearm suicide rates over the study period (Figure A1).

Comment #5: It may also be helpful to add some brief discussion on the difference-in-differences method.

Response #5: We added clarification and additional details about the DID estimation method in the Analysis subsection on pp. 9-11. This includes the additional paragraph about the new "placebo test." We used the canonical two-group, two-period setup. We added the Wing et al. (2018) reference for the reader on p. 9 directly after introducing this approach. It's an intuitive paper and will provide additional technical details for the reader.

Reference:

Wing C, Simon K, Bello-Gomez RA: Designing Difference in Difference Studies: Best Practices for Public Health Policy Research. Annu Rev Public Health; 2018; 39:453–469

Comments from *Reviewer #3*

Comment #1: Great paper, well written.

Response #1: Thank you for your thoughtful review. Your support is much appreciated.

Comment #2: They should add 95% CIs!

Response #2: We have updated Table 2, replacing the parenthetical SEs with 95% confidence intervals. We also added the 95% CIs in the Results section whenever we describe the multivariable estimates.

Comments from *Reviewer #4*

Comment #1: The paper is very well-written and engaging. The study is well presented and motivated.

Response #1: Thank you for your helpful and supportive review!

Comment #2: Can you clarify whether you are using linear regression on suicide rate (#suicides/population size) for each county-year as your outcome of interest? If so, have you considered using Poisson or negative binomial regression instead to account for the varying population sizes across counties? I know this means that the interpretation of the interaction term is then multiplicative but was curious about how/why you made this decision.

Response #2: Yes, we used a linear estimation method, which has performed well for the suicide rate outcomes in previous publications. We considered a GLM approach, but the outcome did not follow an obvious Poisson (or negative binomial) distribution. The distribution was fairly symmetric. There was also evidence of a linear relationship between the firearm suicide rate and time. For example, the correlation between the firearm suicide rate and year for Pima County was 0.78 over the study period. We believe the linear estimation approach is robust and helps the reader interpret the DID estimate of the effect of the 2016 law on firearm suicide rates in Pima County. When we estimate a Poisson regression model, the direction and statistical significance of the estimates are the same as when we used the linear estimation method. For instance, the DID estimate in a Poisson regression model is 0.095 (P < 0.0001), shown in Table B below. Exponentiated, the IRR for the DID estimate = 1.101. We also get results nearly identical to what we present in the paper when we log-transform our dependent variable and use linear estimation methods.

Table B. DID estimates from the adjusted Poisson regression model

	1 *Adjusted Model*
Policy variables	
SB 1487 exposure	
Comparison group	Ref
Policy group (Enactment of state law, SB 1487, preempting gun disposal ordinance in Tucson, Pima County)	0.100**
	(0.011)
Policy enactment timing	
Pre-law enactment	Ref
Post-law enactment	0.034

	(0.027)
Policy group x Post-law enactment (difference-in-differences estimate)	0.096**
	(0.016)

Notes: Excerpted policy variable estimates from the fully-adjusted Poisson regression model. ** indicates P < 0.001

Comment #3: Have you considered including yearly fixed effects in lieu of the post-law enactment period variables, and including a categorical “time_since_policy_change” variable in place of interaction

term (so that this variable equals 0 in all untreated states and in pre policy time in AZ, 1 in the first year after treatment in AZ, then 2 in the second year and so on)? In particular, it might be interesting to see if they effect over time was dynamic. It sort of looks like the effect increases with time.

Response #3: We have used “flexible” DID and two-way fixed-effects (TWFE) approaches in previous papers. However, we do not feel comfortable using that approach for this analysis. First, we are concerned about the recent literature on covariate exogeneity issues and bias (and asymptotic bias) with using TWFE estimators for DID applications (e.g., Wooldridge, Sant-Anna, Imai and Kim). Second, and more importantly, our analysis only uses 54 observations. We are concerned about dimensionality problems that arise from including an additional 13 indicator variables in our models, plus the new mental health professional shortage area indicator. We would not have ample data to power the models using this approach. Given the limitations of this study, we believe a parsimonious two-group, two-period setup is the best first step while explicitly stating the assumptions you recommend below. That said, we agree it would be interesting to see if the policy effects change over time in follow-up studies. We added this point to our *Conclusions* section on p.

17. We also called attention to the increase in Pima County’s unadjusted firearm suicide rate by the end of the study in the *Discussion* section on p. 13.

Comment #4: When you describe the covariates you adjust for, it would be helpful to include a sentence about what sorts of variables are confounders in DID analysis (variables that are state and time varying).

Response #4: This is an excellent recommendation to help the reader understand our approach. We added the following to the *Covariates* subsection on p. 9: “Our empirical approach assumes that confounders varying across the policy and comparison groups are time-invariant and time-varying confounders are group invariant.” We also clarified our adjustment for additional confounders varying across the groups.

Comment #5: Either in the methods or the discussion, would also be nice to see a sentence addressing the third assumption of DID — that there was no other change at the time of the policy change that affected Tucson/Pima county and not the other counties. Need to at least mention it as an assumption and perhaps discuss in discussion if there may be something else happening that you know about over that time period.

Response #5: Thank you for this helpful suggestion. We added this important assumption to the *Analysis* subsection on p. 10: "This approach also assumed that there were no other unmeasured policy changes or factors coinciding with the timing of Arizona's 2016 preemption law that could have affected firearm suicide rates in Pima County relative to the comparison group counties." As recommended, we also discussed this important caveat to our final *Limitations* paragraph on p. 16 about other, unobserved characteristics. As described in our response to Reviewer #2, we also decided to conduct a "placebo test" of the expected policy effects and pre-policy common trends assumption. For this placebo test, we performed an additional DID estimation using a "fake" policy group for our outcome of interest. Specifically, we replicated our estimation of Model 2 using Maricopa County for our policy group observations and all other non-Pima counties for the comparison group. This test and the results of the test are described in the revised manuscript (pp. 10-12).

Comment #6: Round effect estimate and CI to two digits ini abstract and results and tables.

Response #6: All decimal places have been rounded to two digits in the Abstract, Results, and Tables.

Comment #7: Describe the effect estimates as risk differences (increase in additive risk of outcome due to policy change)

Response #7: We added a sentence at the beginning of the *Discussion* section on p. 13 communicating the effect estimate in terms of additive (or relative change in) risk: "Relative to the comparison counties, the 2016 law coincided with a 10.9% relative increase in the firearm suicide rate in Pima County from the pre-policy period to the post-policy period."

Comment #8: Replace p-values with 95% CIs in results

Response #8: P-values have been replaced with 95% CIs.

Comment #9: Though interesting, the info on R-squared about the adjusted models is superfluous - don't need a higher R^2 for causal models, just need a model that adjusts/eliminates confounding through the model specification so I suggest removing this from the results.

Response #9: That is a great point. We removed the R^2 from the *Results* section.

Comment #10: I would also suggest removing the sentence in the results about the association between unemployment and firearm availability and suicide rates since it isn't a result related to your study objectives.

Response #10: We removed the sentence in the *Results* section about the association between unemployment and suicide. If the reviewer is amendable, we would like to keep the sentence about the association between our firearm availability proxy and the outcome, as this relationship is key to the identification of the effect of firearms on suicide (but not non-firearm suicides) and related to a key limitation of this study and others like it in the scholarly literature.

Comment #11: Rather than talking about 4 models - suggest reframing as 2 models - unadjusted and adjusted - for your primary outcome, and then say you ran the same set of models as part of a sensitivity analysis or negative control test.

Response #11: That is an excellent suggestion for organizing and communicating our key findings. We moved our Models 3 and 4 (robustness test) results from Table 2 to supplemental Table A2, updating the *Results* subsection accordingly.

VERSION 2 – REVIEW

REVIEWER	Ku, Benson Emory University School of Medicine
REVIEW RETURNED	22-Mar-2022

GENERAL COMMENTS	All concerns from this reviewer has been addressed.

REVIEWER	Tu, Wei Georgia Southern University College of Science and Mathematics, Geology and Geography
REVIEW RETURNED	27-Mar-2022

GENERAL COMMENTS	I appreciated authors' great efforts in thoroughly addressing the reviewers' questions. I am overall happy with the revised paper. I have some minor questions/suggestions for the authors to consider. In Response #2, the authors argued that "Evaluating the impact of the original ordinance is beyond the scope of our study", I agree. However, the impact of the original ordinance is critical in justifying the overall the logic of the paper. The general logic chain is as follow, some individuals who have decided to end their lives by the means of firearms was unable to do so due to the difficulty in obtaining firearms, the ordinance (allowing firearm auctions) have made the firearm more accessible to this group of individuals and some were able to obtain the firearms directly or indirectly through the auctions and then ended their lives using the purchased firearms. The statistical models have supported the reasoning by showing significantly rising firearm suicide rate in Pima county after the enactment of Arizona's 2016 preemption law. . I still feel that the above reasoning less optimal. Figure 2 showed "the unadjusted firearm suicide rate decreased in Pima County from 10.7 per 100k in 2006 to 8.9 per 100k in 2007 (and 8.8 per 100k in 2008)." The same figure also showed that the unadjusted firearm suicide rate has been on an overall upward trend from 2008 and 2016 for both the control and comparison groups. No, we certainly cannot make any conclusions based on the figure, it seems that if or to what extent the ordinance has helped reduce the availability of firearms and then contributed to lower the firearm suicide rates remains to be elusive. If the effectiveness of the original ordinance is unclear, then the conclusion that "state-level policy efforts to

	preempt or limit local government from enacting firearm safety and control policies – especially policies that decrease the availability of firearms in local communities…" has become less convincing. Another implicit but quite strong assumption of this research is that there is a large group of suicide attempters who would not consider using non firearm means to end their lives but to wait until firearms are available. I suggest the authors add some brief discussions to make overall logic chain as well as the conclusions stronger.

VERSION 2 – AUTHOR RESPONSE

Comments from *Reviewer #2*

Comment #1: I appreciated authors' great efforts in thoroughly addressing the reviewers' questions. I am overall happy with the revised paper. I have some minor questions/suggestions for the authors to consider.

> **Response #1: Thank you for your continued support and thoughtful review. It is much appreciated.**

Comment #2: In Response #2, the authors argued that "Evaluating the impact of the original ordinance is beyond the scope of our study", I agree. However, the impact of the original ordinance is critical in justifying the overall the logic of the paper.

The general logic chain is as follow, some individuals who have decided to end their lives by the means of firearms was unable to do so due to the difficulty in obtaining firearms, the ordinance (allowing firearm auctions) have made the firearm more accessible to this group of individuals and some were able to obtain the firearms directly or indirectly through the auctions and then ended their lives using the purchased firearms. The statistical models have supported the reasoning by showing significantly rising firearm suicide rate in Pima county after the enactment of Arizona's 2016 preemption law. I still feel that the above reasoning less optimal. Figure 2 showed "the unadjusted firearm suicide rate decreased in Pima County from 10.7 per 100k in 2006 to 8.9 per 100k in 2007 (and 8.8 per 100k in 2008)." The same figure also showed that the unadjusted firearm suicide rate has been on an overall upward trend from 2008 and 2016 for both the control and comparison groups. No, we certainly cannot make any conclusions based on the figure, it seems that if or to what extent the ordinance has helped reduce the availability of firearms and then contributed to lower the firearm suicide rates remains to be elusive. If the effectiveness of the original ordinance is unclear, then the conclusion that "state-level policy efforts to preempt or limit local government from enacting firearm safety and control policies – especially policies that decrease the availability of firearms in local communities…" has become less convincing. Another implicit but quite strong assumption of this research is that there is a large group of suicide attempters who would not consider using non firearm means to end their lives but to wait until firearms are available. I suggest the authors add some brief discussions to make overall logic chain as well as the conclusions stronger.

> **Response #2: We made the following revisions to address this helpful suggestion further:**

1. **We revised the *Conclusions* section on p. 16 (marked version) to readdress the logic for the reader. We added your language about how the effects of the original 2005 ordinance remain elusive. However, we also added available evidence about the number of firearms eliminated from the 2005 ordinance. We reiterated the important connection between firearm availability and suicide risk while referencing previous literature that contradicts the "substitution hypothesis." People in suicidal crisis who do**

not have easy access to firearms (highly-lethal suicide methods) do not typically substitute to other ways to attempt suicide (e.g., Daigle, 2005). Firearm availability matters greatly. Notably, we found the 2016 preemption law did not impact non-firearm suicide rates in Pima County relative to other counties over the same period.

2. **We revised the *Limitations* subsection on p. 14 and the *Introduction* section on p. 5 to soften the logic about how SB 1487 "likely" affected firearm availability.**
3. **We revised the *Limitations* subsection to start with the limitation about the potential influence of other unobserved characteristics (i.e., other policies coinciding with SB 1487 that may have affected firearm suicide rates in Pima County during the same period) and data limitations (e.g., not being able to measure actual firearm availability in this or other studies like it). We stressed these points as key limitations for the reader considering our logic. Our findings were not causal.**