

Andersen, Synøve N. et al.

Working Paper

Replicating "State Action to Prevent Violence against Women". A Comment on Córdova & Kras (2022)

I4R Discussion Paper Series, No. 67

Provided in Cooperation with:
The Institute for Replication (I4R)

Suggested Citation: Andersen, Synøve N. et al. (2023) : Replicating "State Action to Prevent Violence against Women". A Comment on Córdova & Kras (2022), I4R Discussion Paper Series, No. 67, Institute for Replication (I4R), s.l.

This Version is available at:
<https://hdl.handle.net/10419/276961>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



No. 67

I4R DISCUSSION PAPER SERIES

Replicating “State Action to Prevent Violence against Women”. A Comment on Córdova & Kras (2022)

Synøve N. Andersen

Nicolai T. Borgen

Solveig T. Borgen

Elisa Brini

Michael J. Frith

Alexi Gugushvili

September 2023

I4R DISCUSSION PAPER SERIES

I4R DP No. 67

Replicating “State Action to Prevent Violence against Women”. A Comment on Córdoba & Kras (2022)

**Synøve N. Andersen¹, Nicolai T. Borgen², Solveig T. Borgen¹,
Elisa Brini¹, Michael J. Frith¹, Alexi Gugushvili¹**

¹University of Oslo/Norway

²Oslo Metropolitan University/Norway

SEPTEMBER 2023

Any opinions in this paper are those of the author(s) and not those of the Institute for Replication (I4R). Research published in this series may include views on policy, but I4R takes no institutional policy positions.

I4R Discussion Papers are research papers of the Institute for Replication which are widely circulated to promote replications and meta-scientific work in the social sciences. Provided in cooperation with EconStor, a service of the [ZBW – Leibniz Information Centre for Economics](#), and [RWI – Leibniz Institute for Economic Research](#), I4R Discussion Papers are among others listed in RePEc (see IDEAS, EconPapers). Complete list of all I4R DPs - downloadable for free at the I4R website.

I4R Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

Editors

Abel Brodeur
University of Ottawa

Anna Dreber
Stockholm School of Economics

Jörg Ankel-Peters
RWI – Leibniz Institute for Economic Research

Replicating “State Action to Prevent Violence against Women”.

A comment on Córdoba & Kras (2022)

Synøve N. Andersen (University of Oslo)¹

Nicolai T. Borgen (Oslo Metropolitan University)

Solveig T. Borgen (University of Oslo)

Elisa Brini (University of Oslo)

Michael J. Frith (University of Oslo)

Alexi Gugushvili (University of Oslo)

Abstract

Córdoba and Kras (2022) examine how the existence of a women’s police station (WPS) in the place of residence influences citizens’ attitudes toward gender-based violence in Brazil. In their analytical specification, the authors find that men are more likely to reject violence against women (VAW) and support bystander intervention in municipalities with a WPS, especially if the WPS has been operating for a long time. This paper examines the replicability and robustness of Córdoba & Kras’ (2022) findings. First, we reproduce the paper’s main findings and uncover one minor coding error and three estimates that have been reported with the opposite sign compared to that in our reproduction; neither is of consequence for the study’s main results. Second, we test the robustness of the results by (1) recoding one of the main explanatory variables and several of the control variables to account for non-linear trends, (2) using alternative techniques to estimate clustered standard errors, (3) consistently applying a 95% confidence level in the presentation of the results, (4) altering the propensity score matching (PSM) procedure as well as the composition of the variables used in the PSM robustness check, (5) using an alternative technique to test for multicollinearity, (6) excluding potential endogenous control variables, and (7) using an alternative coding for computing margins. Reassuringly, the results are robust to most of these tests. However, two of the robustness checks challenge parts of the paper’s main findings. First, allowing for non-linearity in the effect of time since the establishment of WPS shows (a) a non-linear effect on VAW and (b) no apparent changes in either male or female attitudes over time once the WPS has been established. Second, the inclusion of other variables in the PSM procedure renders part of the main estimates of interest statistically nonsignificant ($p < 0.1$). Our findings highlight the need for further re-analyses and replications for investigating the preventive effects of women’s police stations.

¹ Corresponding author. Email: synovena@uio.no. Phone: +4790219003

1 Introduction

This paper presents a replication and robustness check of Córdoba and Kras' (2022) paper on state action to prevent violence against women. In their paper, Córdoba and Kras (2022) examine whether the existence of a women's police station (WPS) in the place of residence influences citizens' attitudes toward gender-based violence in Brazil. Drawing on Brazilian cross-sectional survey data from 2013 matched with municipality-level indicators (pages 5-6), the authors estimate whether WPS causally impacts attitudes toward violence against women (VAW) using multilevel mixed-effects ordered logistic models (page 7). The paper has two main findings. The first main finding is that men are more likely to express intolerance toward VAW (Figure 1 on page 10) and bystander intervention attitudes (Figure 3 on page 11) in municipalities with a WPS. The second main finding is that the longer a WPS has existed in a municipality, the higher are men's probability of rejecting VAW (Figure 2 on page 10) and expressing bystander intervention attitudes (Figure 4 on page 11). The authors conclude that the implementation of women's police stations has the potential to address gender-based violence by changing men's attitudes (page 16).

In this study, we investigate the replicability and robustness of Córdoba and Kras' (2022) main findings. We successfully reproduced the paper's main results; we uncovered one minor coding error and three estimates that have been reported with the wrong signs, as well as one estimate and one standard error that diverged at the third decimal, but neither was of consequence for the study's main conclusions. We then turned to sensitivity analyses. We tested the robustness of the results by (1) recoding one of the main explanatory variables and several of the control variables to account for non-linear trends, (2) using alternative techniques to estimate clustered standard errors, (3) consistently applying a 95% confidence level in the presentation of the results, (4) altering the composition of variables used in the propensity score matching (PSM) robustness check, as well as the PSM procedure, (5) using an alternative technique to test for multicollinearity, (6) excluding potential endogenous control variables, and (7) using an alternative coding to consider interactions. With two major exceptions, the results are robust to these tests.

Two of our robustness checks challenge parts of the paper's main findings. First, relaxing the assumption that time since the establishment of WPS has a linear effect on VAW has implications for one of the paper's main findings. Contrary to the authors' conclusion of stronger effects if the WPS has been present for a long time, this robustness check suggests that time since the establishment of WPS has a non-linear effect on attitudes towards VAW and that after the WPS has been established,

no additional effect is observed over time for either gender. Second, as a sensitivity test, the authors use propensity score matching (PSM) to address endogeneity. However, when we include other variables in the PSM procedure, parts of the main estimates of interest become statistically nonsignificant ($p=0.204$ and $p=0.236$). This highlights the need for further re-analyses and replications for investigating this effect.

In the following, section 2 briefly summarizes the hypothesis, data, methods, and results of Córdoba and Kras (2022). Then, in section 3, we report on the computational reproducibility, using the author's original code and data. In section 4, we report on our various robustness checks before providing a brief conclusion in section 5. The authors' original data and Stata code were downloaded from the online publication site.²

2 A brief summary of Córdoba & Kras (2022)

Córdoba and Kras (2022) examine whether the existence of a women's police station (WPS) in the place of residence influences citizens' attitudes toward gender-based violence in Brazil. They have two main hypotheses (Córdoba & Kras 2022, page 5):

Hypothesis 1: Men who live in towns with a WPS will be more likely to reject VAW and express bystander intervention attitudes than their counterparts in towns without a WPS.

Hypothesis 2: The longer a WPS has been enacted in the place of residence, the higher the men's probability of rejecting VAW and expressing bystander intervention attitudes will be.

To test their hypotheses, they use data from the 2013 Data Popular and the Instituto Patricia Galvão cross-sectional survey, which is nationally representative of the adult population in urban areas, matched with municipal-level indicators from various sources (Córdoba & Kras 2022, pages 5-7). The survey data includes 1,501 face-to-face interviews carried out in 100 randomly selected municipalities across the five regions in Brazil, where respondents are asked about their degree of intolerance toward

² Available at <https://journals.sagepub.com/doi/full/10.1177/0010414019879959>

VAW and bystander intervention attitudes. These survey responses are used as dependent variables and are measured using a Likert scale from 1 to 5, where 1 means “strongly disagree” and 5 “strongly agree”.

The 2013-survey data is complemented with municipal-level indicators from the Institute of Geography and Statistics, the National Secretariat of Public Safety, Mapa da Violência, Mapa de Coletivo de Mulheres, Google News archives, Igarapé Institute, and UNDP. At the municipal level, the core independent variables are whether the municipality has established a WPS (coded 1 if established, 0 otherwise) and the number of years since the creation of the first WPS in the municipality. The authors report that in the sample, 46% of municipalities have at least one WPS (a few have more than one WPS), and the age of WPS across municipalities ranges from 2 to 28 years, with municipalities without a WPS coded as 0.

The paper estimates the effects of WPS on VAW using multilevel mixed-effects ordered logistic models with random effects for the intercept at the municipal level (Córdova & Kras 2022, page 7). The ordered logistic model is used because the dependent variables are ordinal, while the two-level multilevel model is used to account for the nested structure of the data. Interaction terms are included between gender and WPS and between gender and years since WPS creation to test hypotheses 1 and 2, respectively. The paper uses observed control variables to account for municipal and individual-level confounding.

The paper has two main findings. The first is that men are more likely to express intolerance toward VAW (Córdova & Kras, 2022: Figure 1, page 10) and bystander intervention attitudes (Córdova & Kras, 2022: Figure 3, page 11) in municipalities with WPS, supporting hypothesis 1. Specifically, as shown in Figure 1, their regression models predict that 63.8% of men living in municipalities without a WPS strongly agree that women should not be beaten, while in municipalities with a WPS, the share who strongly agree with this claim is 80.6%. For women, about 82% strongly agree that women should not be beaten, both in municipalities with and without WPS. Concerning bystander intervention attitudes, 81.7% of men living in a municipality with a WPS strongly agree that a person should call the police if witnessing VAW, compared to only 56.0% of men living in a municipality without a WPS. There is also a difference among women, at about 80% vs. 70%, respectively, but this difference is not statistically significant at the 5% level.

The second main finding is that the longer a WPS has existed in a municipality, the higher are men’s probability of rejecting VAW (Córdova & Kras, 2022: Figure 2, page 10) and expressing bystander

intervention attitudes (Córdova & Kras, 2022: Figure 4, page 11), supporting hypothesis 2. Men who reside in municipalities where WPS was created 28 years ago are about 20 percentage points more likely to strongly agree that women should not be beaten than men residing in municipalities without WPS. Likewise, the authors report a similar change in bystander intervention attitudes, with men who are residing in the oldest WPS municipalities being more than 20 percentage points more likely to strongly agree that a person should call the police if witnessing VAW than men residing in municipalities without WPS.

3 Reproducibility

To test the computational reproducibility of the results, we rerun the analysis in Stata 17 using the do-files and data files provided by the authors. The do-file runs without error messages, but there is one typo that overwrites Figure 1 with Figure 8. When we change this by replacing “chart1” with “chart8” on line 360 of the original do-file, the reproduced graphs are identical to those included in the paper.

The results in Table 1 in Córdova and Kras (2022) are identical to the reproduced output except for five estimates and one standard error. Three estimates have the opposite sign as in the manuscript and two estimates and one standard error diverge in the third decimal:

- Model 1: The SE for *Femicide rate* is 0.044 in the reproduced output and 0.043 in the manuscript.
- Model 3: The estimate for *Wealth* is 0.015 in the reproduced output and -0.015 in the manuscript.
- Model 4: The estimate for *Wealth* is 0.039 in the reproduced output and -0.039 in the manuscript.
- Model 7: The estimate for *Age 35-50* is 0.074 in the reproduced output and 0.073 in the manuscript.
- Model 8: The estimate for *Higher education* is -1.057 in the reproduced output and 1.057 in the manuscript.
- Model 8: The estimate for *No. years since WPS creation* is 0.049 in the reproduced output and 0.048 in the manuscript.

The reproduced Table 1 is included in Appendix Table A1. All cells with diverging values are shaded grey.

4 Robustness

4.1 Recoding of independent variables

In the following two sections, we test to what extent the results are robust to alternative specifications of some of the main independent variables. We start with one of the key explanatory variables, namely “years since establishment of WPS”, we then move on to some of the controls.

4.1.1 *Non-linear transformations of “years since establishment of WPS”*

The authors assume that men’s attitudes toward gender-based violence are linearly improved as a function of the number of years since establishing a WPS. Models 4 and 8 in the original manuscript include “years since establishment of WPS” as a continuous variable with a linear specification (Córdova and Kras, 2022, page 8). Value 0 on this time variable represents municipalities without a WPS, while 2, 4, and 6 represent WPSs that have existed for 2, 4, and 6 years, respectively, and so on. We test the robustness of the main results to non-linear transformations of the variable “years since establishment of WPS” using the same data and the same model as the original study.

We alter the original operationalization of the “years since establishment of WPS” variable by including several non-linear transformations that relax the assumption that time since establishment of WPSs has a linear impact on the likelihood of rejecting VAW and the likelihood of supporting bystander intervention. More specifically, we first rerun models 4 and 8 with the original linear specification of the time variable before running models with quadratic, cubic, and quartic transformations of the time variable. We also run a model with a non-parametric specification of the time variable. The set of dummies included in the non-parametric specification has the value 0 (no WPS) as the reference category, and dummies indicating whether the WPS was established 1-4, 5-9, 10-14, 15-19, 20-24, or 25-28 years prior to the survey.

The top row in Figure 1 shows the reproduced results included in Figure 2 in the original study (i.e., the results from analyses with a linear specification of the time variable). Panel A shows the effect of time since establishment on attitudes towards VAW separately for men and women, whereas Panel B shows how the gender gap has changed with time since WPS establishment. From these panels, the original study concludes that men are more likely to reject VAW in municipalities with WPS, especially if WPS has been operating for a long time (see Córdova and Kras, 2022, pages 7 and 10).

The next six panels show results from models including quadratic, cubic, and quartic specifications of the time variable. The panels can be compared vertically: all panels in the same column include results from models that are identical to those of Córdova and Kras (2022) – except for the operationalization of the explanatory variable time since WPS establishment. The results reveal a non-linear effect of the time since establishment of WPS on intolerance to VAW. When including a quartic term (Panels G and H), the results show that establishment of WPS closes the gender gap in the likelihood of rejecting VAW, which is in accordance with results from the original model. However, there is no evidence that men are more likely to oppose VAM if WPSs have been operating for a long time. According to Panel G, there are no apparent changes over time in either male or female attitudes. Further, Panel H also shows that there is no decrease in the gender gap over time once WPS is established, and the main difference in intolerance towards VAW thus seems to be between municipalities with and without WPS. The two bottom-most panels of Figure 1 show results using a non-parametric specification of the time variable using dummy variables. There is no clear pattern over time in this model, but it should be noted that the results from the models using the dummy specification are rather imprecise.

[Figure 1]

Figure 2 shows replicated results of the effect of time since the establishment of WPS on bystander interventional attitudes. The top panels reproduce Figure 4 from the original study. Based on these panels, the original study concludes that men are more likely to support bystander interventions in municipalities where the WPS has been operating for a long time (see Córdova and Kras, 2022, pages 10-11).

The lower panels of Figure 2 again show results from models including quadratic, cubic, and quartic specifications of the time since establishment variable. The quartic specification (Panels G and H) reveals an increase in the likelihood of agreeing with bystander interventions during the first period after establishment, but then the likelihood remains more or less stable over time. As with intolerance toward VAW, the gender gap in attitudes towards bystander interventions remains statistically nonsignificant after the establishment of WPS. Thus, we do not find support for the claim that men are more likely to support bystander interventions if the WPS has been operating for a long time, and there is little evidence that the change in attitudes is stronger for men than women. It is also worth mentioning that panel I, with the non-parametric specification of the time variable, shows a dip in attitudes toward bystander intervention in the first 1-4 years after the establishment of WPS. This dip is particularly pronounced for men, and the size of the estimate is larger than the long-term increase

in supportive attitudes for this group. While the estimated attitude level after 1-4 years is imprecise, this pattern warrants a discussion of potential mechanisms that might go against the authors' expectations.

[Figure 2]

4.1.2 *Control variables*

In the original paper, four ordinal independent variables with just five levels are treated as continuous variables: “VAW private issue”, “approves traditional gender roles”, “perceives gender discrimination”, and “social class”. Although treating ordinal variables as continuous is both justifiable and commonly done, some have argued for extra caution when there are few levels in the ordinal variable. For example, Johnson and Creech (1983) suggest that this is important for variables with up to four levels, while others suggest that one should test the impact of treating ordinal variables as continuous irrespective of the number of levels (Long & Freese, 2006). Based on these recommendations, we assess the robustness of the overall results by treating these four variables as a series of dummy-coded binary indicator variables rather than continuous variables. The results for the main variables of interest when using this approach are presented in Table 1. As shown, there is little substantive difference between the two sets of results in terms of the coefficient size, standard errors, or significance level for the main variables of interest. The only variable affected in significance is the number of years since the WPS creation in the final model (8), which falls from significant at $p < .1$ level to statistically nonsignificant. We therefore find the (main) results to be largely robust to this change in specification.

[Table 1]

4.2 **Clustering of standard errors**

In the paper, the authors deal with the clustering of the 1,501 survey responses within the 100 municipalities by estimating a multilevel model, which we will call multilevel-clustering. However, the clustering issue could also be accounted for by using a multilevel model with cluster-robust standard errors. An advantage of this clustering approach is that it also allows for heteroskedasticity (e.g., Huang et al., 2022). As such, we re-estimate the main models from the paper (models 1a-8a in Table 1) using multilevel-clustered standard errors and then using a multilevel model with cluster-robust standard

errors. As the clustering will not affect coefficients, only the difference in standard errors and the changes to significance level are presented in Table 2.

As shown in Table 2, there are few substantive differences in the standard errors and significance levels when computing the multilevel cluster-robust standard errors compared to multilevel-clustering standard errors. The standard errors change on average (regardless of direction) by 0.03, which is a small change relative to the size of the standard errors (see Table 1). Also, while there are 23 (of 152 possible) changes in significance levels, only six involve changing to or from not statistically significant, given an alpha level of 0.10. For example, this applies to the differences in bystander intervention attitudes between men and women in models 5 and 7 – which the authors already concede by describing them as “weak” (Córdova and Kras, 2022, page 7). Adopting cluster-robust standard errors also changes the significance of two other variables though these are included as control variables and are of no substantive interest. One final change is in the standard error of the “perceives gender discrimination” estimate, which changes so that the estimate is significant at the 0.01 level only. As this is still below the typical alpha level of 0.05, this change is inconsequential.

[Table 2]

4.3 Consistently applying a 95% significance level

When presenting the results from the PSM models, Córdova and Kras (2022) opted to show 90% confidence intervals for the estimates in Panel B of **Figure 1**

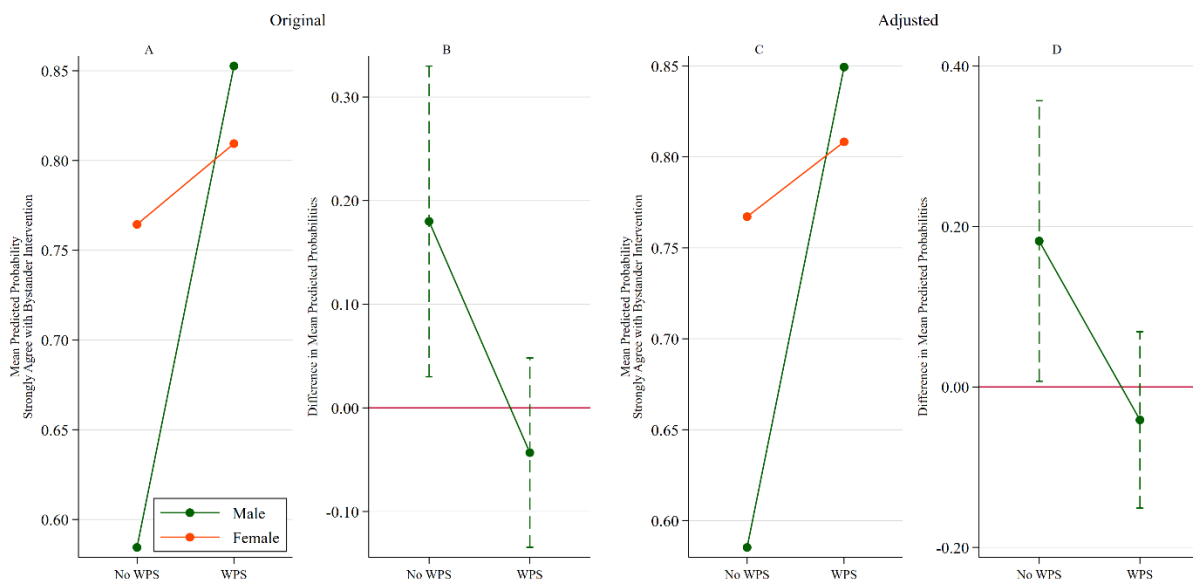


Figure 3 rather than the 95% confidence intervals used for other estimates in the paper. We therefore replicate the results from Figure 3, Panel B, in their manuscript using a 95% confidence level (see Figure 3).

[Figure 3]

Since the 95% confidence interval for the gender difference in panel B does not overlap with 0, using 95% confidence intervals rather than 90% does not change the authors' conclusions. However, the lower bound of the 95% confidence interval would be very close to 0, as shown in

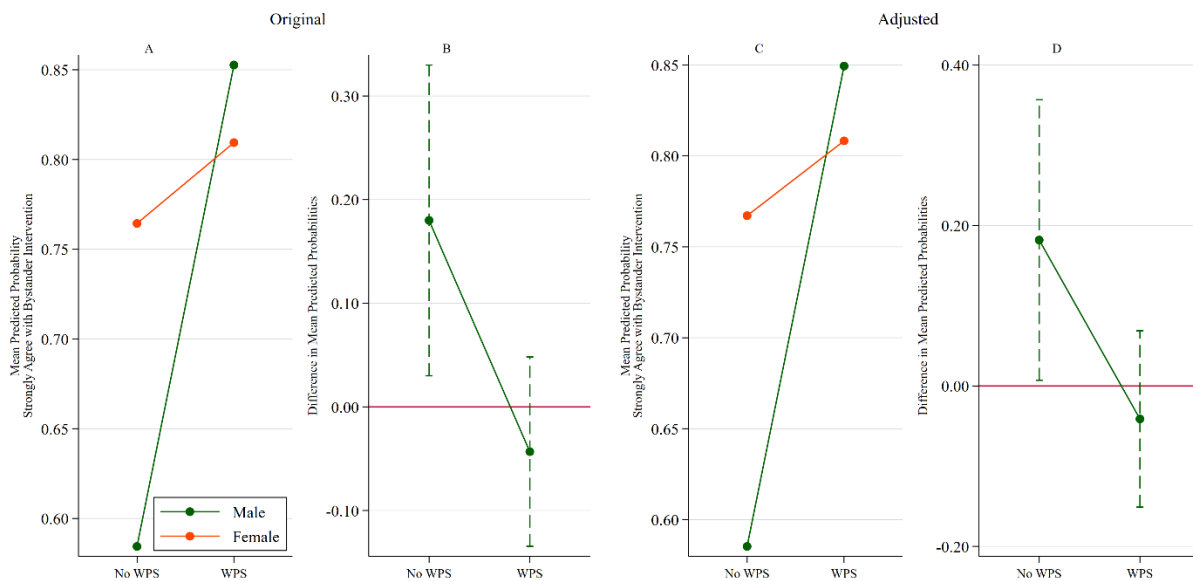


Figure 3(b).

4.4 Propensity score matching

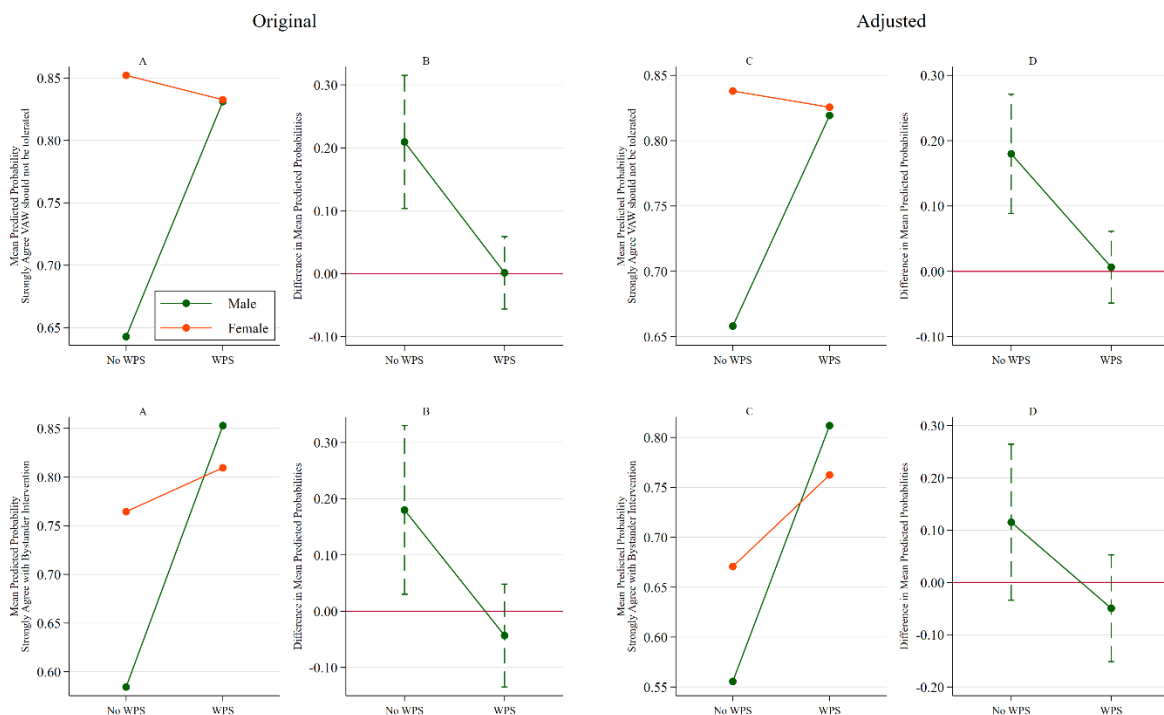
In their paper, Córdova and Kras (2022) aim to mitigate potential endogeneity problems in treatment assignment, i.e., having a WPS in a municipality, by applying a weighting method based on propensity score to match municipalities with and without WPS. The authors select three variables to predict treatment assignment: population size, GDP per capita, and homicide levels. Table 3 shows the authors' actual estimates of propensity score using binary logistic regression, and the homicide rate is not a significant predictor of the WPS establishment. The authors use the variables for 2009-2013 to predict the treatment assignment that predominantly happened before 2009. According to the provided data, for 90% of municipalities which had WPS by 2009-2013, these WPS were established before 2006.

[Table 3]

In addition to the issue of timing of contextual variables used in the PSM, the authors selected the homicide variable for this part of the analysis. The dataset also includes the rate of femicide in the municipalities. Crime that specifically targets women can be a more important predictor of the WPS establishment than the general level of crime measured by the homicide rate. To check the robustness of the original matching, we have substituted the homicide rate variable with the femicide rate variable and estimated the propensity score. The results of this exercise are shown in Table 4.

[Table 4]

We see that the femicide rate, just like the homicide rate, is not a significant predictor of treatment assignment. It is important to note, however, that the femicide rate variable is derived after the period in which many WPSs were established. We moreover see that if we use the femicide rate instead of the homicide rate to generate the propensity score that is used to match the treated and the untreated municipalities, the original estimates of the difference in bystander intervention attitudes between men and women become statistically nonsignificant – even at the 10% level ($p=0.204$).³



³ Please note that we kept the paper’s original 90% significance level for the analysis of bystander intervention attitudes.

Figure 4 shows the estimates in the original article as well as our adjusted estimates where the femicide rate is used instead of the homicide rate.

[Figure 4]

We have also checked how sensitive the results are if we substitute other contextual variables in the propensity score generation procedure. For instance, the authors, in their sensitivity analysis, imply that the level of education can be considered as an alternative measure of the municipalities' development. Based on this reasoning, we substituted the log GDP variable with the mean level of education in the municipalities and then fitted the same specification as the authors did in the original analysis using the “psmatch2” command. The results of this exercise are shown in Table 5 and Figure 5. We can see that the level of education is not a significant predictor of the treatment, along with the homicide rate. More importantly, the original estimates on the gender difference in bystander intervention attitudes are cut in half and again become statistically nonsignificant at the 10% level when these two variables are interchanged ($p=0.236$).

[Table 5]

[Figure 5]

To check the validity of the specific PSM technique used in the analysis, we finally tested the robustness of the results using alternative matching techniques. First, instead of the nearest neighbor with caliper procedure employed by the authors, we used the Kernel matching technique. The results were comparable to the original analysis in this instance. Further, we varied the number of matches of the treatment to control groups to 3, 4, 6, and 7 matches, testing the robustness of the 1 to 5 match used in the original study. Results show that 4, 6, and 7 matches performed similarly to 5 matches, but that using 3 matches to generate the propensity score leaves the gender difference in bystander intervention attitudes statistically nonsignificant at the 10% level ($p=0.105$). Finally, we used an alternative caliper to find suitable matches by adjusting the caliper distance from 0.25 to 0.20, as recommended in the relevant literature (Austin, 2011). This adjustment did not change the main results. Overall, although results are sensitive to the choice of control variables, our sensitivity analysis with several alternative matching techniques did not substantively change the results.

4.5 Multicollinearity test

As noted in Footnote 6 in Córdoba & Kras (2022, page 14), the authors checked for multicollinearity using the common approach of checking the variance inflation factor (VIF) for each variable and concluded that multicollinearity is not a problem in their models. This process often involves estimating a separate OLS regression (here with clustered standard errors) using the variables from the original statistical model(s) and then computing the VIF scores from that OLS regression.

However, the OLS regressions in this process should (ideally) be computed using the same variables in the same form as they were included in the other analyses (here, in the ordinal logistic regressions). In Córdoba and Kras (2022), the education level variable is included as a set of binary indicator variables in the main regression models but treated as a continuous variable in the OLS regressions used for checking the VIF scores (unlike, for example, social class, which is consistently included as binary indicators across all the models). We test the robustness of the multicollinearity test by including education as binary indicator variables; both the original VIF scores and our adjusted VIF scores are presented in Table 6. Notably, the adjusted output includes four new VIF scores per model as we vary the baseline level (which was ‘no schooling’ in the original analysis) for the education level variable (cf. Allison, 2012).

[Table 6]

The first thing to note from Table 6 is that the VIF scores for all variables, except education level, are remarkably consistent regardless of whether education level is included as a continuous variable (baseline = NA) as originally computed or included as a series of dummy variables (baseline = NS, PE, SE, and HE) and regardless of which education level is the baseline level. The second thing to note, though, is the dramatic increase in the VIF score(s) associated with the education level variables when the statistic is calculated from an OLS regression as specified as per the ordinal logistic regression in the original paper (baseline = NS). On the face of it, using the general rule of thumb that values above 10 indicate a potentially problematic level of multicollinearity, the scores of 10-17 across the models would suggest a problem.

At the same time, many argue (e.g., Allison, 2012) that VIF scores related to dummy variables demarcating levels of the same categorical variable are to be cautiously interpreted. This is largely because they will be inevitably multicollinear with each other as having a value in one variable (e.g., having no schooling) will automatically be negatively associated with having a value in the others (e.g.,

primary school education). Furthermore, although the VIF scores computed relate to the precise specifications of the models in the original paper (i.e., using no schooling as the baseline level), others note that the choice of baseline category and the proportion of the sample in that category can have a substantial impact (e.g., Hendrickx et al., 2004). As such, changing the baseline level from no schooling (which has just 28 participants in the full sample), the VIF scores become satisfactory and between 1 and 3. As such, although the VIF scores were originally calculated using a specification of the education variable that did not match that used in the main models, the adjusted VIF scores are consistently satisfactory and in line with Córdova & Kras' (2022) claim that there are no multicollinearity problems in the models.⁴

4.6 Exclusion of endogenous controls

Córdova and Kras (2022) include control variables in their models that might be affected by the implementation and/or time since implementation of a WPS. More specifically, the femicide rate, respondents' perceptions of VAW as a private matter, respondents' opinions about whether there is discrimination against women in society, respondents' support for traditional gender roles, respondents' perceptions of the idea that the state supports victims to denounce, and respondents' approval of female (and leftist) presidents, are all variables that may be affected by whether a WPS is present in their municipality, and how long it has been there. Conditioning on such variables may control away some of the causal effects they are interested in (overcontrol bias). Moreover, if there is unobserved confounding of these endogenous variables, then conditioning on them might introduce collider biases into the estimated effects of WPS on attitudes toward VAW. In keeping with this, we considered alternative model specifications that exclude these variables. Results yield almost identical estimates for the effect of WPS and time since WPS establishment on intolerance toward VAW (Figures 6–7) and bystander intervention attitudes (Figures 8–9).

[Figure 6]

[Figure 7]

[Figure 8]

⁴ We could also vary the baseline level for the other dummy coded variables (age, marital status, and race) for the aforementioned reasons this is arguably not necessary in that they will provide useful information regarding multicollinearity.

[Figure 9]

4.7 Alternative coding for margins

Córdova and Kras (2022) examine whether the existence of a WPS in the place of residence and the length of time a WPS has been operating in a municipality affect Brazilian women’s or men’s likelihood to reject VAW. To test a potential gender gradient, the authors present the average gender gap in the mean predicted probability of rejecting VAW in municipalities with and without WPS and municipalities with different WPS ages.

To estimate predicted probabilities of rejecting VAW separately for men and women, the authors use the Stata command `-margins-`. The Stata command `-margins-` has two options: `-over(x)-` and `-at(x)-`. Using `-over(x)-`, the command loops over the values of variable `x` and computes the specified margin only for observations where `x` takes the current value while ignoring the rest of the data. In other words, for each value of the variable ‘female’, it computes the specified margin using only those observations in the data for which the variable ‘female’ takes on the current value in the loop (0=men; 1=women).

Using the alternative option `-at(x)-`, the command also loops over the specified values of the variable ‘`x`’, but without restricting to the data where ‘female’ takes the corresponding value. It rather creates a virtual data set in which every observation has the variable ‘female’ at that value and computes the margin in this virtual data set. Therefore, the output from the `-at()`- option is a set of margins that are adjusted to the distributions of the covariates in the entire sample (StataCorp, 2021).

In the following figures (Figures 10–13), we compare the author’s estimates (labeled “Original estimates”) with results based on this alternative coding (labeled “Adjusted estimates”). All in all, the substantive interpretation remains unchanged after alternative coding specification, yet some differences are present.

Adjusted estimates in Figure 10 (panels C and D) are in line with the authors’ results (panels A and B) in indicating that in municipalities without a WPS, there is a significant gender gap in the probability of rejecting VAW, while the gender gap closes and is no longer statistically significant in municipalities with a WPS. We do, however, observe that using the alternative `-margins-` option means that there is a smaller gender gap in both municipalities without and with a WPS and that the effect of WPS is somewhat reduced.

[Figure 10]

Figure 11 replicates the effects of WPS's age on intolerance toward VAW. Again, no substantive differences with the authors' findings are detected after changing the coding, at the same time as the gender differences are smaller than those presented in the original paper.

[Figure 11]

Figures 12 and 13 test the authors' findings concerning bystander intervention attitudes. The comparison between Panel A and Panel C in Figure 12 again shows that the alternative -margins- option leads to a smaller gender difference in bystander intervention attitudes both in municipalities with and without a WPS. The lower gender gap is mainly driven by lower predicted probabilities that women strongly agree with bystander intervention. The effect of WPS is also reduced somewhat in the adjusted estimates (panel D) compared to the original estimates (panel B). A Wald-test confirms that gender differences between municipalities with no WPS and with a WPS are still statistically significant at the 1% level.

[Figure 12]

Our alternative coding specification also shows the effect of WPS's age on support for bystander intervention being smaller than the effect reported by the authors (Figure 13, panel D). However, the general pattern remains unchanged.

[Figure 13]

5 Conclusion

In this paper, we have examined the replicability and robustness of Córdova & Kras' (2022) findings. The original paper hypothesizes and finds support for two claims. First, men who live in municipalities with a woman's police station (WPS) are more likely to reject violence against women (VAW) and express bystander intervention attitudes than their counterparts in municipalities without a WPS. Second, the longer a WPS has been operational in the municipality of residence, the higher the men's probability of rejecting VAW and expressing bystander intervention attitudes are. Regarding the second hypothesis, the original study found that men who reside in municipalities where WPS was

created 28 years ago are about 20 percentage points more likely to strongly agree that women should not be beaten than men residing in municipalities without WPS. Likewise, the authors report a similar change in bystander intervention attitudes, with men residing in the oldest WPS municipalities being about more than 20 percentage points more likely than men residing in municipalities without a WPS to strongly agree that a person should call the police if witnessing VAW. As a concluding remark, Córdova & Kras (2022, page 16) claim that WPSs can only result in substantial improvements in men's attitudes toward VAW if WPS are permanent.

We conducted a series of replication exercises to assess the replicability and robustness of the original study's findings. Córdova & Kras' (2022) main conclusions remained unaffected by accounting for minor coding errors (section 3), considering alternative techniques for manipulating control variables (4.1.2), adopting an alternative clustering approach for standard errors (4.2), consistently applying 95% levels of significance throughout the paper (4.3), testing for multicollinearity (4.5), removing potential endogenous controls (4.6), and considering alternative coding for predicting attitudes (4.7). Nevertheless, two robustness tests provide results that warrant further investigation regarding the core conclusions drawn in the original research.

First, our robustness checks of the linearity assumption imply a rejection rather than a confirmation of Córdova & Kras' (2022) second hypothesis. We ran models that allow for non-linearity in the effect of time since the establishment of WPS on the two outcomes. The results show that there are no apparent changes in either male or female attitudes towards VAW over time once the WPS has been established. Further, while there seems to be an increase in the likelihood of agreeing with bystander interventions during the first period after establishment, the likelihood then remains more or less stable over time. Thus, in sum, we find no evidence to support the claim that men are more likely to support bystander interventions if the WPS has been operating for a long time. We also find little evidence that the change in attitudes is stronger for men than women. Thus, while the implementation of women's police stations may have the potential to address gender-based violence by changing men's attitudes, as claimed by Córdova & Kras (2022, page 16), we find no evidence of increasing effect over time. Nor do we find support for the claim that WPSs need to be permanent to result in substantial improvements in men's attitudes (page 16).

Second, our robustness checks of the PSM analyses call for a moderation of Córdova & Kras' (2022) claim that the observed lack of a gender gap in attitudes in municipalities with a WPS suggests that WPS generates policy feedback effects (p. 3 and 14) – at least for bystander intervention attitudes. If

the lack of a gender gap reflects policy feedback effects, then such a gender gap should be present in municipalities without a WPS. However, our robustness test of the propensity score matching procedure shows a statistically nonsignificant gender gap in bystander intervention attitudes in municipalities without a WPS. More specifically, we substituted the homicide rate variable with the femicide rate variable (Figure 4) and the log GDP variable with the level of education variable (Figure 5) in the estimation of the propensity scores. Both approaches lead to results that show no statistically significant difference in bystander intervention attitudes between men and women in the absence of WPS, and no significant differences in the presence of WPS. This result leads to some uncertainty as to the support for Córdova & Kras' (2022) first hypothesis.

In sum, our reproduction and replication of Córdova & Kras' (2022) call into question whether WPS influences attitudes toward gender-based violence and highlight the need for further re-analyses and replications for investigating the preventive effects of women's police stations.

Conflict of interest

The authors have no conflict of interest to declare.

Financial support

None.

References

- Allison, P. (2012). When can you safely ignore multicollinearity. *Statistical horizons*, 5(1), 1-2.
- Austin, P. C. (2011). An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate behavioral research*, 46(3), 399-424.
- Córdova, A., & Kras, H. (2022). State action to prevent violence against women: The effect of women's police stations on men's attitudes toward gender-based violence. *The Journal of Politics*, 84(1), 1-17.
- Hendrickx, J., Pelzer, B., Te Grotenhuis, M., & Lammers, J. (2004). Collinearity involving ordered and unordered categorical variables. RC33 conference in Amsterdam,
- Huang, F. L., Wiedermann, W., & Zhang, B. (2022). Accounting for heteroskedasticity resulting from between-group differences in multilevel models. *Multivariate behavioral research*, 1-21.
- Johnson, D. R., & Creech, J. C. (1983). Ordinal measures in multiple indicator models: A simulation study of categorization error. *American Sociological Review*, 398-407.
- Long, J. S., & Freese, J. (2006). *Regression models for categorical dependent variables using Stata* (Vol. 7). Stata press.
- StataCorp. (2021). *Stata 17 Base Reference Manual*. Stata Press.

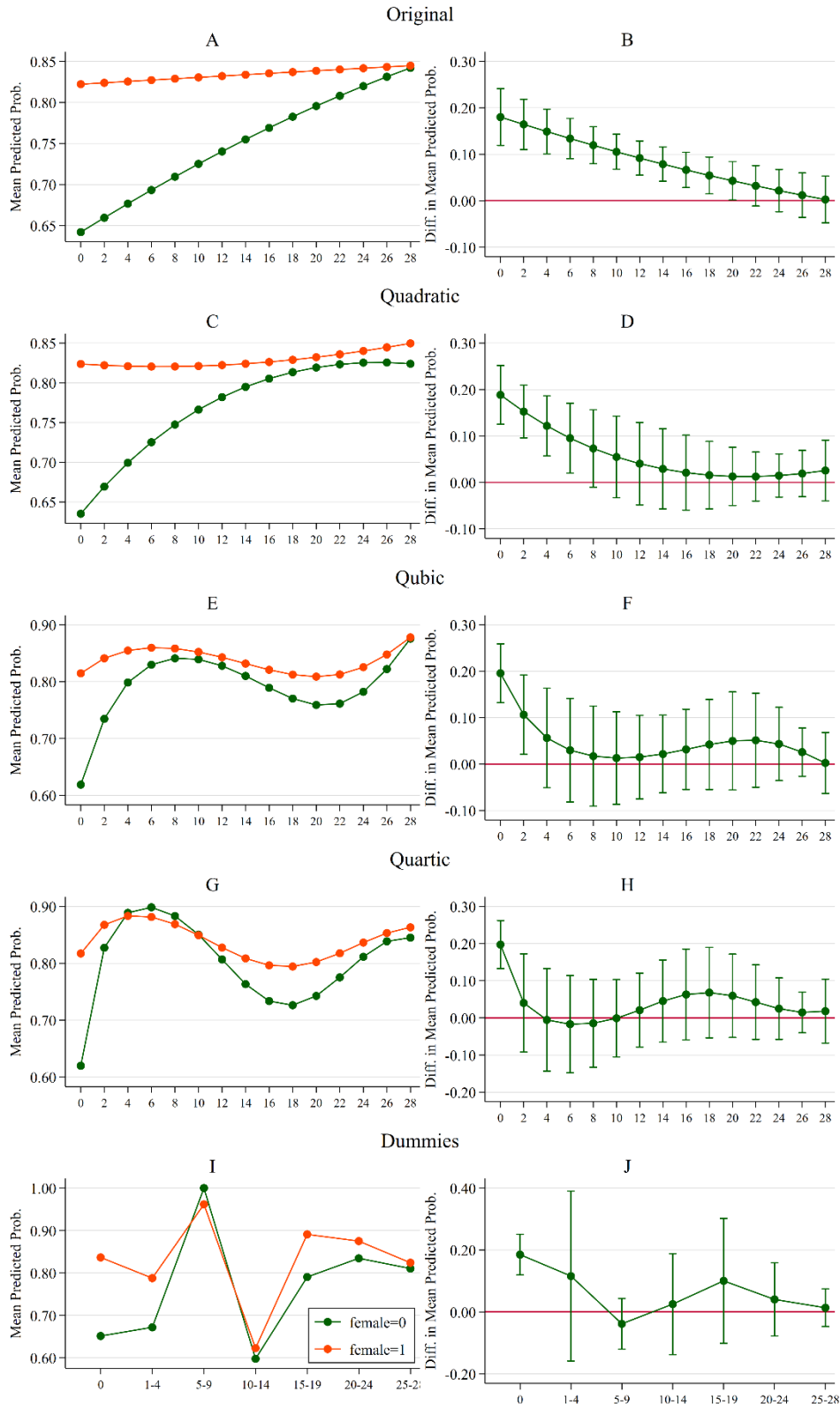


Figure 1: Effect of WPS on intolerance toward VAW. 95% confidence intervals.

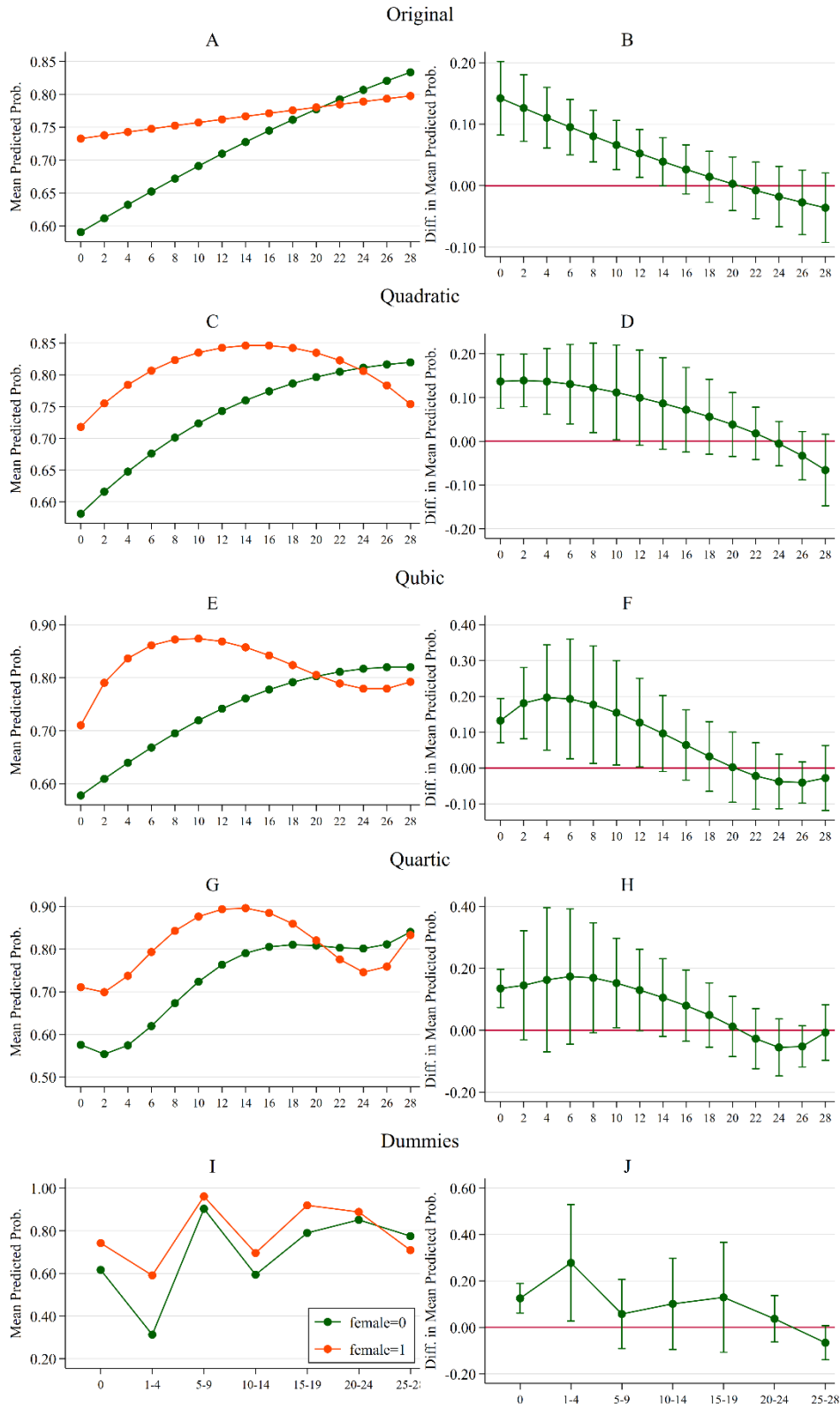


Figure 2: Effect of WPS on bystander intervention attitudes. 95% confidence intervals.

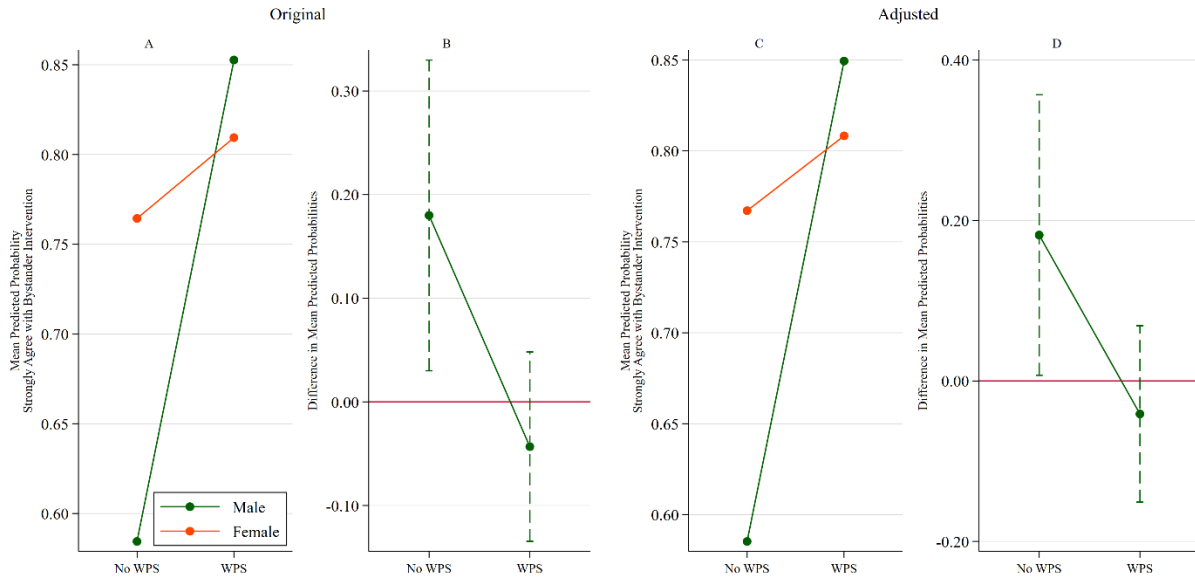


Figure 3: Adjusting the original results by applying the 95% confidence intervals for the Panel B estimates instead of the 90% confidence intervals as used in the original analysis.

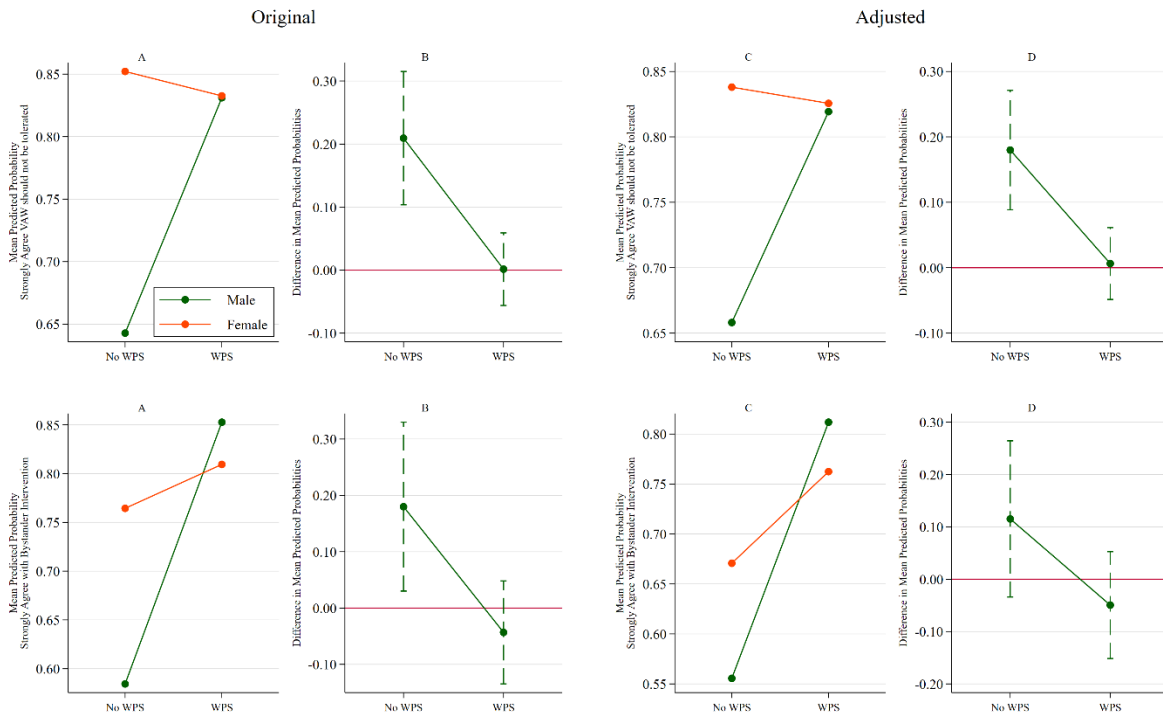


Figure 4: Adjusting the original results by using the femicide rate instead of the homicide rate to match treated and untreated municipalities. For bystander intervention attitudes, 90% confidence intervals are shown.

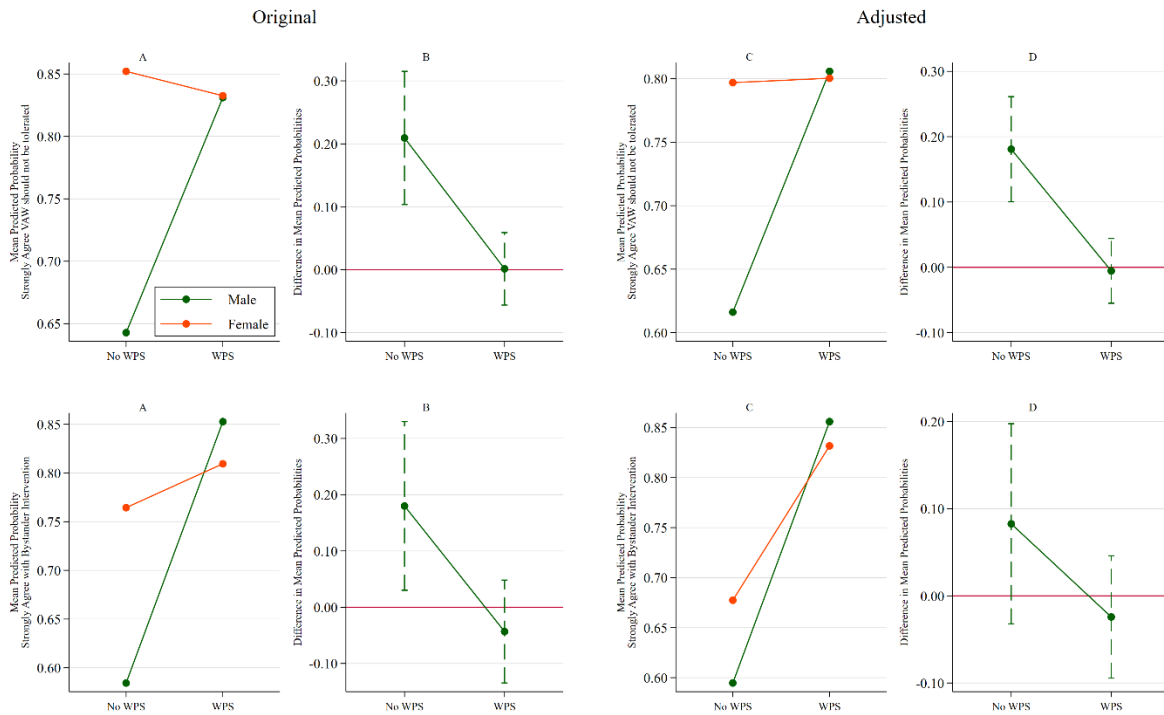


Figure 5: Adjusting the original results by using the level of education instead of the log GDP variable to match treated and untreated municipalities. For bystander intervention attitudes, 90% confidence intervals are shown.

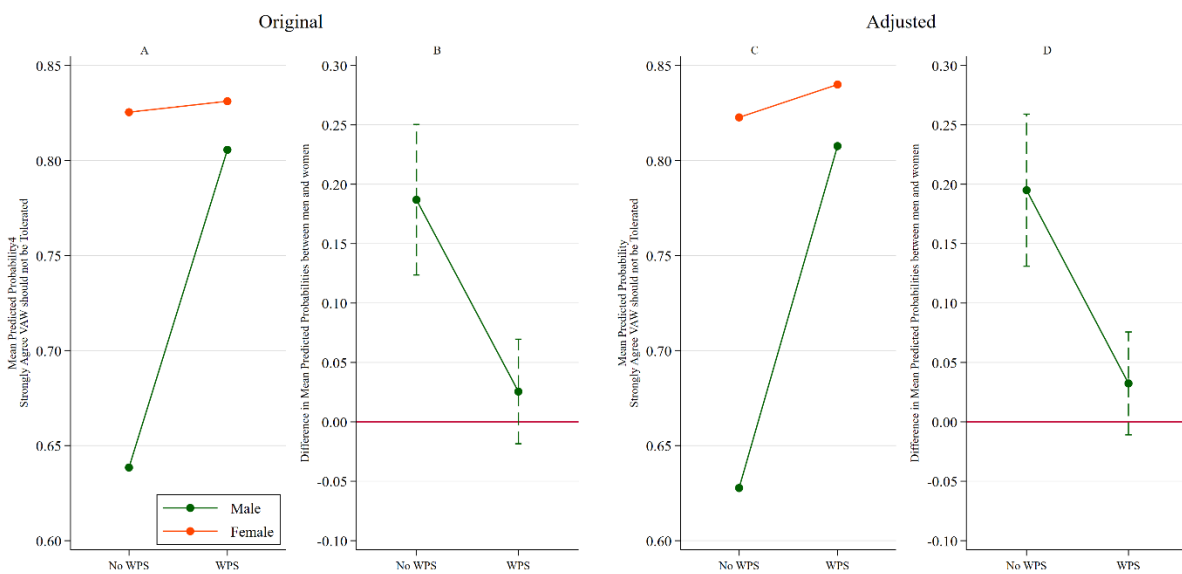


Figure 6: Effect of WPS on intolerance toward VAW. 95% confidence intervals.

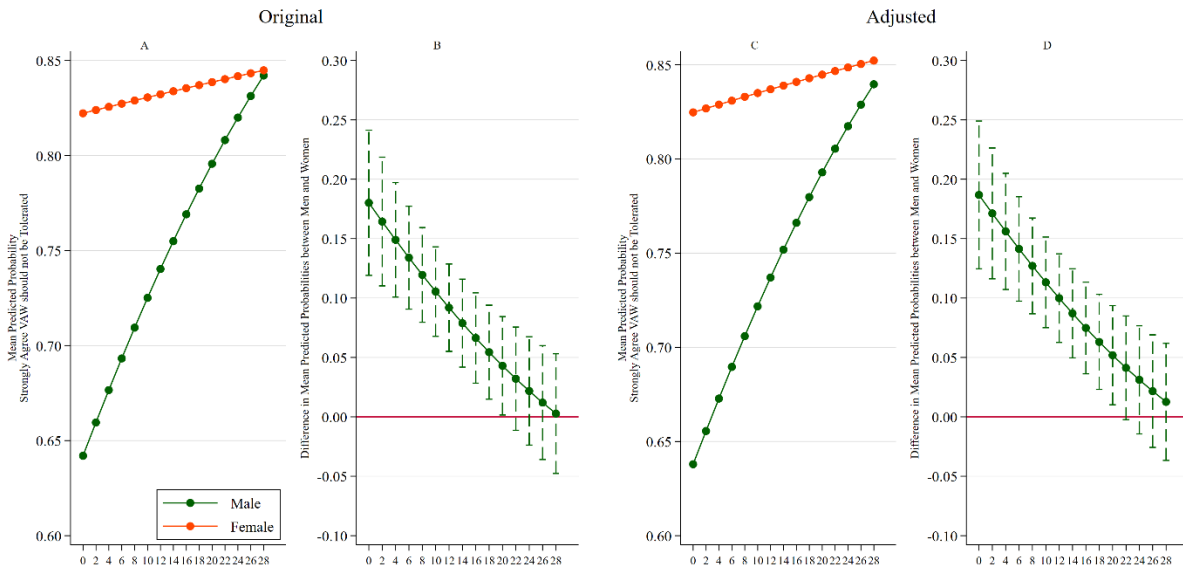


Figure 7: Effect of WPS's age on intolerance toward VAW. 95% confidence intervals.

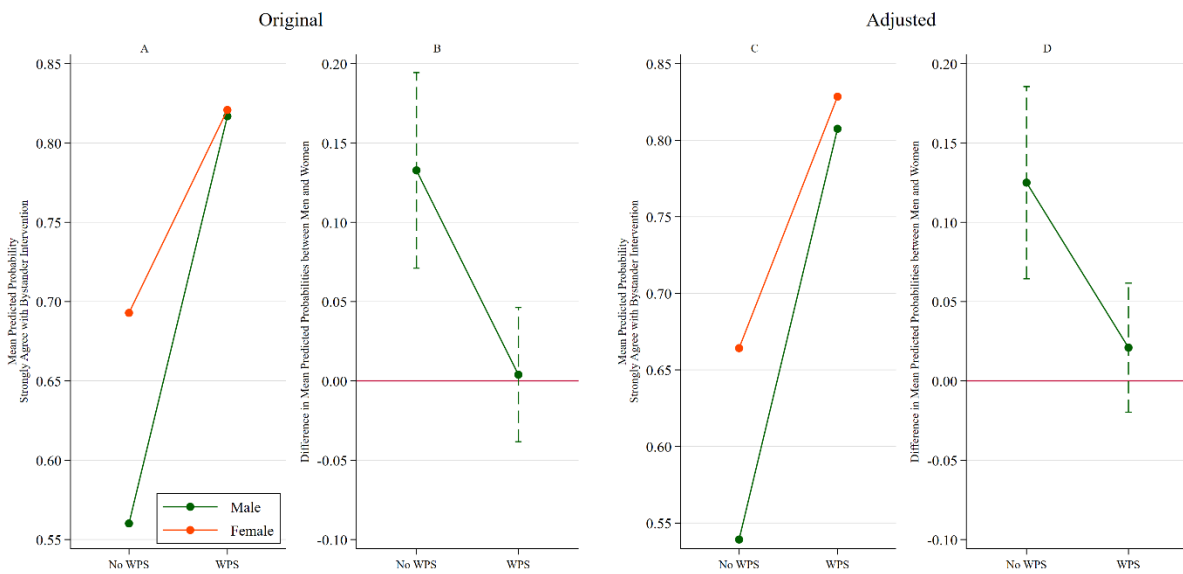


Figure 8. Effect of WPS on bystander intervention attitudes. 95% confidence intervals.

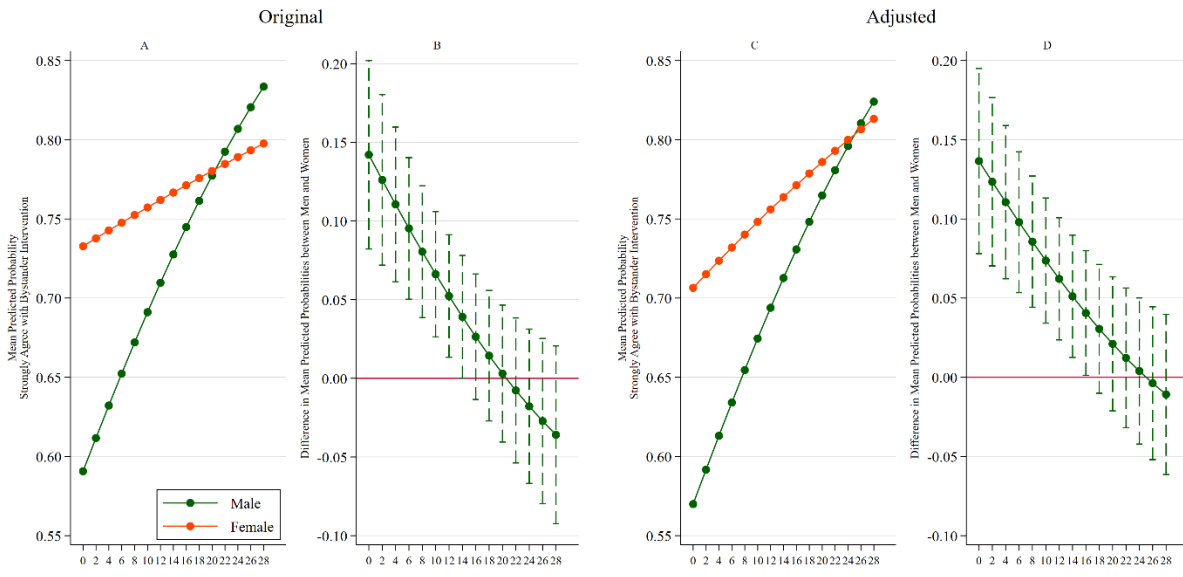


Figure 9: Effect of WPS's age on bystander intervention attitudes. 95% confidence intervals.

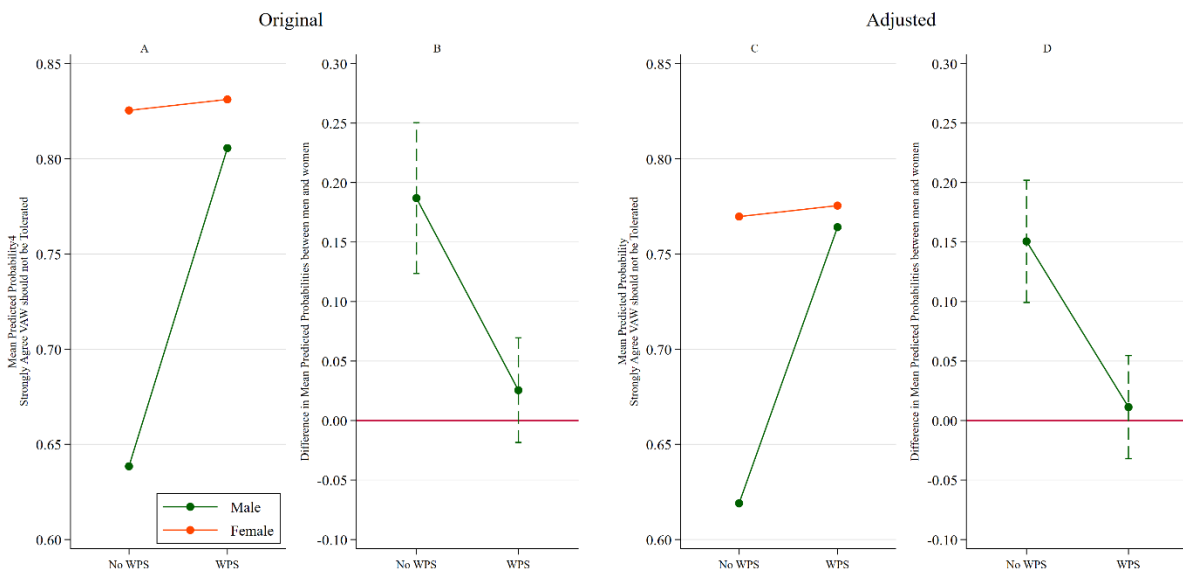


Figure 10. Effect of WPS on intolerance toward VAW using alternative coding for margins. 95% confidence intervals.

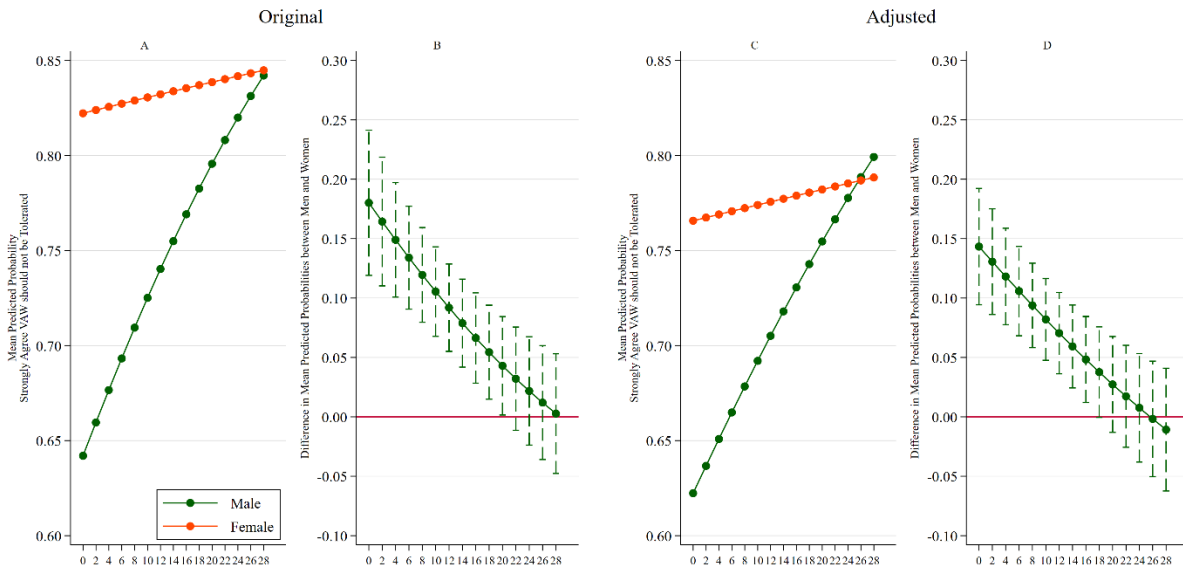


Figure 11: Effect of WPS's age on intolerance toward VAW using alteriative coding for margins. 95% confidence intervals.

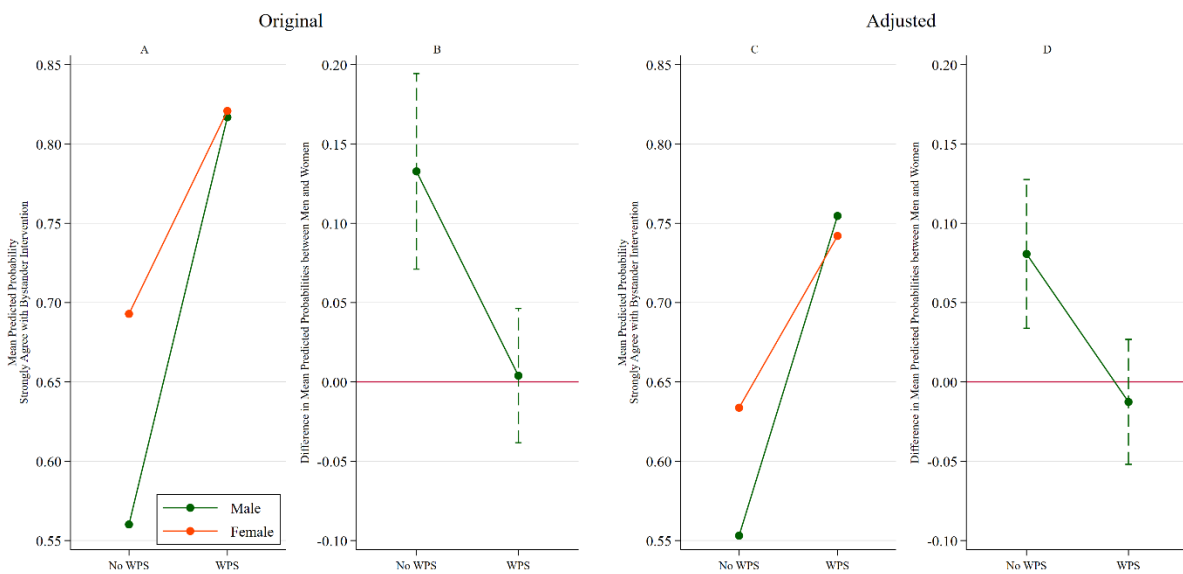


Figure 12: Effect of WPS on bystander intervention attitudes using alteriative coding for margins. 95% confidence intervals.

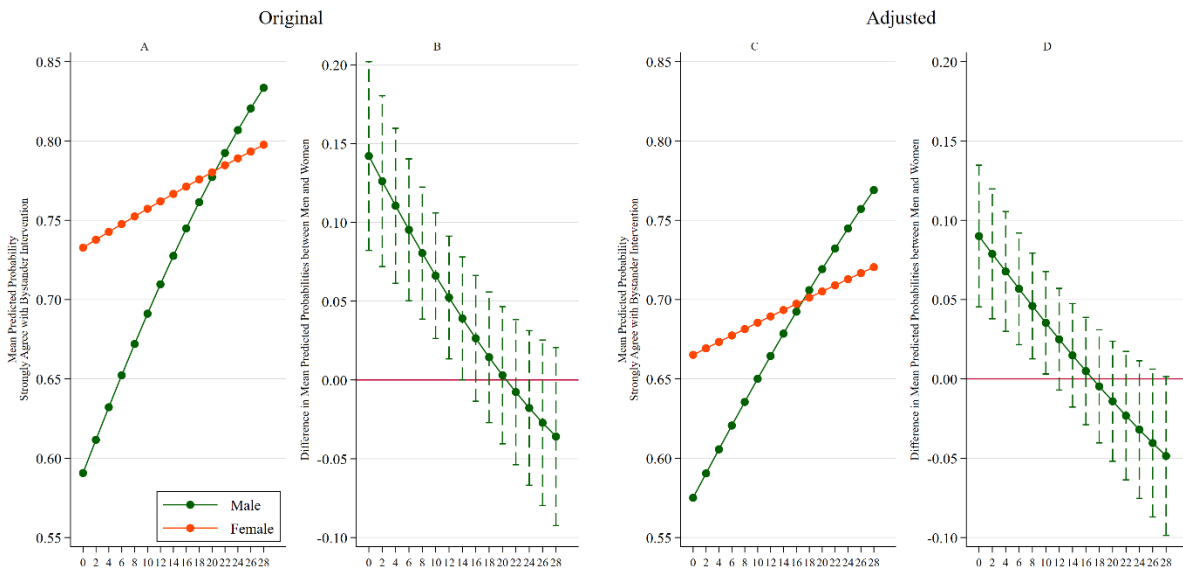


Figure 13: Effect of WPS’s age on bystander intervention attitudes using alteriantive coding for margins. 95% confidence intervals.

Table 1: Determination of attitudes toward violence against women when treating the four five-level variables as continuous (a models) as per the original specification by the authors and when treating them using a series of binary indicator variables (b models)

	Intolerance toward VAW							
	1a	1b	2a	2b	3a	3b	4a	4b
Female (=1; male = 0)	0.535*** (0.139)	0.483** (0.141)	1.041*** (0.202)	0.994*** (0.205)	0.558*** (0.140)	0.504*** (0.142)	0.997*** (0.192)	0.942*** (0.195)
WPS (yes = 1; no = 0)	0.544 (0.466)	0.406 (0.429)	0.927+ (0.485)	0.791+ (0.451)				
WPS x female			-0.884** (0.273)	-0.898** (0.277)				
No. years WPS creation					0.027 (0.023)	0.019 (0.021)	0.042+ (0.024)	0.035 (0.022)
No. years WPS x female							-0.036** (0.012)	-0.036** (0.012)
VAW legislation (yes = 1; no = 0)	0.251 (0.489)	0.370 (0.451)	0.597 (0.547)	0.694 (0.516)	0.276 (0.493)	0.386 (0.455)	0.622 (0.548)	0.701 (0.516)
VAW legislation x female			-0.743 (0.467)	-0.676 (0.475)			-0.762 (0.467)	-0.677 (0.476)
	Bystander Intervention Attitudes							
	5a	5b	6a	6b	7a	7b	8a	8b
Female (=1; male = 0)	0.223+ (0.133)	0.230+ (0.135)	0.530** (0.188)	0.553** (0.191)	0.225+ (0.133)	0.231+ (0.135)	0.600*** (0.182)	0.612*** (0.184)
WPS (yes = 1; no = 0)	1.092+ (0.592)	0.973+ (0.559)	1.396* (0.606)	1.283* (0.576)				
WPS x female			-0.621* (0.261)	-0.625* (0.264)				
No. years WPS creation					0.031 (0.028)	0.025 (0.026)	0.048+ (0.029)	0.042 (0.027)
No. years WPS x female							-0.034** (0.011)	-0.034** (0.011)
VAW legislation (yes = 1; no = 0)	0.009 (0.594)	-0.003 (0.560)	0.035 (0.627)	0.095 (0.597)	0.077 (0.588)	0.055 (0.555)	0.113 (0.622)	0.165 (0.593)
VAW legislation x female			-0.080 (0.412)	-0.218 (0.419)			-0.125 (0.411)	-0.259 (0.418)

Note. Two-level ordered logistic multilevel models. Models control for various variables (see Córdova and Kras, 2022). Standard errors in parentheses.

+ p < .10. * p < .05. ** p < .01. *** p < .001.

Table 2: The difference in standard errors and significance levels when using cluster-robust standard errors in the determination of attitudes toward violence against women

	Intolerance toward VAW				Bystander Intervention Attitudes			
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
Female (=1; male = 0)	0.01	0.05	0.01	0.03	0.04 (+ to NS)	0.00	0.04 (+ to NS)	0.00 (***)
WPS (yes = 1; no = 0)	0.01	0.01			-0.04 (+ to *)	-0.05		
WPS x female	0.00	0.01			0.00	0.09(* to +)		
No. years WPS creation			0.00	0.00			0.00	0.00
No. years WPS x female				0.00				0.00(** to *)
VAW legislation (yes = 1; no = 0)	0.12	0.02	0.13	0.05	-0.04	-0.02	-0.03	-0.03
VAW legislation x female		-0.11(NS to *)	0.00	-0.11(NS to *)		0.05	0.00	0.01
Municipal-level control variable								
Regular police (yes = 1; no = 0)	-0.12	-0.12	-0.09	-0.09	-0.11	-0.11	-0.09	-0.09
Femicide rate	0.00	0.00	0.00	0.00	0.01	0.01	0.01	0.01
Log GDP per capita	-0.02(* to **)	-0.02(* to **)	-0.02	-0.02(* to **)	-0.01	-0.01	0.00	0.00
Log population size	-0.01	-0.01	0.00	0.01	-0.02(NS to +)	-0.02(NS to +)	-0.02	-0.02
Individual-level control variable								
VAW private issue	0.01	0.01	0.01	0.01	0.00	0.00	0.00	0.00
Approves traditional gender roles	0.01	0.01	0.01	0.01	0.02	0.02	0.02	0.02
Perceives gender discrimination	0.01(** to *)	0.01(** to *)	0.01(** to *)	0.01(** to *)	0.03(** to *)	0.03(** to *)	0.03(** to *)	0.04(***)
Perceives state supports victims	0.01	0.02	0.01	0.02(** to *)	-0.01	-0.01	-0.01	-0.01
Government approval	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Primary education (=1; no schooling = 0)	0.08	0.07	0.09	0.08	-0.06	-0.06	-0.06	-0.07
Secondary education	0.10	0.09	0.11	0.11	-0.06	-0.06(NS to +)	-0.06	-0.06
Higher education	0.12	0.11	0.13	0.13	-0.06	-0.05	-0.06	-0.06
Wealth	0.00	-0.01	-0.01	-0.01	0.03	0.03	0.03	0.03
Age 35-50 (=1; 18-34 = 0)	-0.01	-0.01	-0.01	-0.01	0.01	0.01	0.01	0.01
Age 50+	0.01	0.01	0.01	0.01	0.00	0.00	0.00	0.00

Note: Two-level ordered logistic multilevel models. Models control for race, marital status, number of children, and perceived class. The figures outside and inside parentheses, respectively, are the difference in standard errors from using multilevel-clustering standard errors to using a multilevel model with cluster-robust standard errors and the change, if relevant, in significance level. NS $p \geq 0.10$, + $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$.

Table 3: Predictors of treatment assignment using “psmatch2” command

	Coefficient	Std. errors	z	P>z	[95% conf. interval]	
Homicide rate	-0.007	0.009	-0.840	0.399	-0.025	0.010
Log GDP	1.121	0.418	2.680	0.007	0.302	1.939
Log population size	1.046	0.246	4.250	0.000	0.564	1.529
Intercept	-22.79	4.418	-5.160	0.000	-31.445	-14.13

Table 4: Predictors of treatment assignment using “psmatch2” command, the homicide rate is substituted with the femicide rate

	Coefficient	Std. errors	z	P>z	[95% conf. interval]	
Femicide rate	-0.013	0.069	-0.190	0.847	-0.127	0.100
Log GDP	1.225	0.403	3.040	0.002	0.562	1.888
Log population size	0.943	0.199	4.750	0.000	0.616	1.270
Intercept	-22.79	4.490	-5.080	0.000	-30.18	-15.41

Table 5: Predictors of treatment assignment using “psmatch2” command, the log GDP variable is substituted with the level of education variable

	Coefficient	Std. errors	z	P>z	[95% conf. interval]	
Homicide rate	-0.014	0.009	-1.580	0.114	-0.029	0.001
Log population size	1.225	0.301	4.060	0.000	0.729	1.720
Level of education	2.408	2.865	0.840	0.401	-2.304	7.120
Intercept	-15.32	2.869	-5.340	0.000	-20.04	-10.60

Table 6: Variance inflation factors (VIFs) for different models when treating education level as a continuous variable (baseline=NA) and as a series of dummy variables (baseline=NS, PE, SE, and HE) and when using different education levels as the baseline level

Baseline	Intolerance toward VAW					Bystander intervention attitudes					Intolerance toward VAW					Bystander intervention attitudes					
	NA	NS	PE	SE	HE	NA	NS	PE	SE	HE	NA	NS	PE	SE	HE	NA	NS	PE	SE	HE	
WPS	3.53	3.53	3.53	3.53	3.53	3.51	3.51	3.51	3.51	3.51											
Years since WPS creation											4.60	4.61	4.61	4.61	4.61	4.59	4.60	4.60	4.60	4.60	4.60
Female	2.20	2.20	2.20	2.20	2.20	2.19	2.20	2.20	2.20	2.20	1.98	1.99	1.99	1.99	1.99	1.98	1.98	1.98	1.98	1.98	1.98
WPS x Female	3.11	3.11	3.11	3.11	3.11	3.10	3.10	3.10	3.10	3.10											
Years since WPS creation x Female											2.87	2.87	2.87	2.87	2.87	2.86	2.86	2.86	2.86	2.86	2.86
VAW Legislation	2.16	2.17	2.17	2.17	2.17	2.15	2.16	2.16	2.16	2.16	2.17	2.17	2.17	2.17	2.17	2.16	2.16	2.16	2.16	2.16	2.16
VAW Legs. X Female	2.24	2.25	2.25	2.25	2.25	2.23	2.24	2.24	2.24	2.24	2.25	2.25	2.25	2.25	2.25	2.23	2.24	2.24	2.24	2.24	2.24
Regular police	1.19	1.20	1.20	1.20	1.20	1.19	1.20	1.20	1.20	1.20	1.21	1.22	1.22	1.22	1.22	1.21	1.22	1.22	1.22	1.22	1.22
Femicide rate	1.15	1.15	1.15	1.15	1.15	1.15	1.15	1.15	1.15	1.15	1.15	1.15	1.15	1.15	1.15	1.15	1.15	1.15	1.15	1.15	1.15
Log GDP per capita	1.90	1.90	1.90	1.90	1.90	1.90	1.90	1.90	1.90	1.90	1.84	1.84	1.84	1.84	1.84	1.84	1.84	1.84	1.84	1.84	1.84
Log Pop Size	2.53	2.56	2.56	2.56	2.56	2.52	2.55	2.55	2.55	2.55	3.62	3.65	3.65	3.65	3.65	3.61	3.64	3.64	3.64	3.64	3.64
VAW Private Issue	1.09	1.10	1.10	1.10	1.10	1.09	1.10	1.10	1.10	1.10	1.14	1.14	1.14	1.14	1.14	1.14	1.14	1.14	1.14	1.14	1.14
Approves Traditional gender roles	1.17	1.17	1.17	1.17	1.17	1.17	1.17	1.17	1.17	1.17	1.18	1.18	1.18	1.18	1.18	1.18	1.18	1.18	1.18	1.18	1.18
Perceives gender discrimination	1.05	1.05	1.05	1.05	1.05	1.05	1.05	1.05	1.05	1.05	1.06	1.06	1.06	1.06	1.06	1.05	1.06	1.06	1.06	1.06	1.06
Perc. State supports victims	1.04	1.04	1.04	1.04	1.04	1.04	1.04	1.04	1.04	1.04	1.04	1.04	1.04	1.04	1.04	1.04	1.04	1.04	1.04	1.04	1.04
Government approval	1.16	1.16	1.16	1.16	1.16	1.16	1.16	1.16	1.16	1.16	1.16	1.16	1.16	1.16	1.16	1.16	1.16	1.16	1.16	1.16	1.16
Level education (continuous)	1.44					1.44					1.46					1.46					
No Schooling		BL	1.09	1.14	1.26		BL	1.09	1.14	1.26		BL	1.09	1.14	1.26		BL	1.09	1.14	1.26	
Primary education		17.06	BL	1.41	2.90		17.03	BL	1.41	2.89		16.80	BL	1.42	2.94		16.78	BL	1.42	2.93	
Secondary education		15.60	1.23	BL	2.37		15.57	1.23	BL	2.37		15.41	1.24	BL	2.38		15.38	1.24	BL	2.38	

Higher education	10.69	1.57	1.47	BL		10.73	1.58	1.47	BL		10.47	1.58	1.46	BL		10.51	1.59	1.47	BL	
Social class	1.38	1.38	1.38	1.38	1.38	1.38	1.39	1.39	1.39	1.39	1.38	1.38	1.38	1.38	1.38	1.38	1.38	1.38	1.38	1.38
Wealth	1.39	1.41	1.41	1.41	1.41	1.40	1.41	1.41	1.41	1.41	1.39	1.40	1.40	1.40	1.40	1.40	1.41	1.41	1.41	1.41
Age 35-50	1.47	1.48	1.48	1.48	1.48	1.48	1.48	1.48	1.48	1.48	1.47	1.47	1.47	1.47	1.47	1.47	1.48	1.48	1.48	1.48
Age 50+	1.92	1.93	1.93	1.93	1.93	1.92	1.94	1.94	1.94	1.94	1.91	1.93	1.93	1.93	1.93	1.91	1.93	1.93	1.93	1.93
Single with partner	1.62	1.62	1.62	1.62	1.62	1.61	1.61	1.61	1.61	1.61	1.63	1.63	1.63	1.63	1.63	1.63	1.63	1.63	1.63	1.63
Married	2.74	2.75	2.75	2.75	2.75	2.75	2.75	2.75	2.75	2.75	2.75	2.75	2.75	2.75	2.75	2.75	2.76	2.76	2.76	2.76
Civil union	1.78	1.78	1.78	1.78	1.78	1.79	1.79	1.79	1.79	1.79	1.79	1.79	1.79	1.79	1.79	1.79	1.79	1.79	1.79	1.79
Separated/Divorced	1.56	1.56	1.56	1.56	1.56	1.56	1.56	1.56	1.56	1.56	1.56	1.56	1.56	1.56	1.56	1.56	1.56	1.56	1.56	1.56
Widow	1.58	1.58	1.58	1.58	1.58	1.57	1.57	1.57	1.57	1.57	1.59	1.59	1.59	1.59	1.59	1.58	1.58	1.58	1.58	1.58
Children	1.54	1.55	1.55	1.55	1.55	1.54	1.55	1.55	1.55	1.55	1.54	1.55	1.55	1.55	1.55	1.53	1.55	1.55	1.55	1.55
Black	1.26	1.26	1.26	1.26	1.26	1.26	1.26	1.26	1.26	1.26	1.26	1.27	1.27	1.27	1.27	1.26	1.27	1.27	1.27	1.27
Parda	1.28	1.28	1.28	1.28	1.28	1.28	1.28	1.28	1.28	1.28	1.29	1.29	1.29	1.29	1.29	1.29	1.29	1.29	1.29	1.29
Amarela	1.05	1.05	1.05	1.05	1.05	1.05	1.05	1.05	1.05	1.05	1.05	1.05	1.05	1.05	1.05	1.05	1.05	1.05	1.05	1.05
Indigenous	1.04	1.05	1.05	1.05	1.05	1.04	1.05	1.05	1.05	1.05	1.04	1.05	1.05	1.05	1.05	1.04	1.05	1.05	1.05	1.05
Mean VIF	1.68	2.92	1.65	1.65	1.74	1.67	2.92	1.65	1.65	1.73	1.74	2.96	1.71	1.71	1.79	1.73	2.96	1.71	1.71	1.79

Note: BL denotes the baseline level, Baselines: NA=Not applicable (education level treated as continuous), NS=No schooling, PE=Primary education, SE=Secondary education, HE=Higher education

Appendix – Reproduction of the main table (Table1) from the original study.

Table A1: Reproduced Table 1 from Córdova & Kras (2022). All cells where the values diverge between the reproduction and the original paper are shaded in grey.

	Intolerance towards VAW				Bystander Intervention Attitudes			
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
Female (= 1; male = 0)	0.535*** (0.139)	1.041*** (0.202)	0.558*** (0.140)	0.997*** (0.192)	0.223+ (0.133)	0.530** (0.188)	0.225+ (0.133)	0.600*** (0.182)
WPS (yes = 1; n = 0)	0.544 (0.466)	0.927+ (0.485)			1.092+ (0.592)	1.396* (0.606)		
WPS x female		-0.884** (0.273)				-0.621* (0.261)		
No. Years WPS creation			0.027 (0.023)	0.042+ (0.024)			0.031 (0.028)	0.049+ (0.029)
No. Years WPS x female				-0.036** (0.012)				-0.0345** (0.0108)
VAW legislation (yes = 1; no = 0)	0.251 (0.489)	0.597 (0.547)	0.276 (0.493)	0.622 (0.548)	0.009 (0.594)	0.035 (0.627)	0.077 (0.588)	0.113 (0.622)
WAV legislation x female		-0.743 (0.467)		-0.762 (0.467)		-0.080 (0.412)		-0.125 (0.411)
Municipal-level control variable:								
Regular police (yes = 1; no = 0)	0.417 (0.796)	0.422 (0.806)	0.486 (0.809)	0.494 (0.816)	0.945 (0.990)	0.956 (0.990)	0.871 (0.986)	0.887 (0.988)
Femicide rate	-0.053 (0.044)	-0.054 (0.044)	-0.053 (0.044)	-0.054 (0.044)	-0.059 (0.055)	-0.060 (0.055)	-0.056 (0.055)	-0.056 (0.055)
Log GDP per capita	-0.734* (0.296)	-0.752* (0.300)	-0.682* (0.288)	-0.696* (0.290)	0.230 (0.369)	0.229 (0.369)	0.297 (0.355)	0.298 (0.356)
Log population size	0.116 (0.133)	0.123 (0.135)	0.074 (0.159)	0.081 (0.160)	-0.275 (0.169)	-0.275 (0.169)	-0.209 (0.192)	-0.210 (0.193)
Individual-level control variable:								
VAW private issue	-0.019 (0.055)	-0.034 (0.056)	-0.016 (0.056)	-0.035 (0.056)	-0.133* (0.055)	-0.138* (0.055)	-0.137* (0.055)	-0.147** (0.055)
Approves traditional gender roles	-0.232*** (0.058)	-0.233*** (0.059)	-0.231*** (0.059)	-0.229*** (0.059)	-0.039 (0.060)	-0.035 (0.060)	-0.034 (0.060)	-0.029 (0.060)
Perceives gender discrimination	0.201** (0.072)	0.206** (0.072)	0.205** (0.073)	0.205** (0.073)	0.236** (0.074)	0.243** (0.074)	0.238** (0.074)	0.245*** (0.074)
Perceives state supports victims	0.375* (0.133)	0.404* (0.135)	0.391* (0.159)	0.422** (0.160)	0.520*** (0.169)	0.548*** (0.169)	0.520*** (0.192)	0.558*** (0.193)

	(0.156)	(0.157)	(0.158)	(0.159)	(0.152)	(0.153)	(0.152)	(0.153)
Government approval	0.032	0.029	0.035	0.0319	-0.013	-0.013	-0.013	-0.014
	(0.029)	(0.030)	(0.030)	(0.030)	(0.032)	(0.032)	(0.032)	(0.032)
Primary education (= 1; no schooling = 0)	-0.211	-0.255	-0.226	-0.264	-0.604	-0.594	-0.604	-0.616
	(0.542)	(0.543)	(0.544)	(0.545)	(0.561)	(0.564)	(0.561)	(0.564)
Secondary education	-0.315	-0.363	-0.333	-0.373	-0.956+	-0.941	-0.956+	-0.966+
	(0.555)	(0.556)	(0.557)	(0.558)	(0.576)	(0.579)	(0.576)	(0.579)
Higher education	-0.485	-0.510	-0.486	-0.514	-1.065+	-1.031+	-1.069+	-1.057+
	(0.588)	(0.589)	(0.591)	(0.592)	(0.603)	(0.606)	(0.603)	(0.606)
Wealth	-0.041	-0.012	0.015	0.039	-0.033	-0.026	-0.021	-0.018
	(0.174)	(0.175)	(0.176)	(0.177)	(0.174)	(0.175)	(0.174)	(0.175)
Age 35-50 (= 1; 18-34 = 0)	-0.125	-0.124	-0.131	-0.133	0.075	0.081	0.074	0.067
	(0.176)	(0.177)	(0.178)	(0.179)	(0.167)	(0.168)	(0.167)	(0.168)
Age 50+	-0.215	-0.193	-0.231	-0.223	-0.194	-0.163	-0.192	-0.165
	(0.209)	(0.211)	(0.211)	(0.213)	(0.204)	(0.205)	(0.204)	(0.205)
No. observations	1,358	1,358	1,337	1,337	1,356	1,356	1,335	1,335
No. municipalities	100	100	98	98	100	100	98	98

Note. Two-level ordered logistic multilevel models. Models control for race, marital status, number of children, and perceived class. The effects of these variables are not statistically significant (results not shown). Standard errors in parentheses. + $p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.