Replication of a Research Claim from Weidmann & Callen (2013),
from *the British Journal of Political Science*

Replication Team: Eric L. Sevigny and Jared Greathouse

Research Scientist: Andrew Tyner

Action Editor: Kevin Esterling

Independent Reviewer:

Reviewer #1: Landon Schnabel

Review Period: June 24 - June 29

View-only links to: Original Paper, Original Materials, Replication Materials, Replication Analyses

# Replication of a Research Claim from Weidmann & Callen (2013), from British Journal of Political Science

SCORE report Weidmann_BritJournPoliSci_2013_JRpA_mk67

Sourcing notes for this preregistration are here.

Privacy Statement: Other teams are making predictions about the outcomes of many different studies, not knowing which studies have been selected for replication. As a consequence, the success of this project requires full confidentiality of the research process, including peer review. This includes privacy about which studies have been selected for replication and all aspects of the discussion about these replication designs.

## Table of contents

9. Randomization

Sampling Plan

10. Existing data
11. Explanation of existing data
12. Data collection procedures
13. Sample size
15. Stopping rule

Variables

16. Manipulated variables
17. Measured variables
18. Indices

Analysis Plan

19. Statistical models
20. Transformations
21. Inference criteria
22. Data exclusion
23. Missing data
24. Exploratory analysis
25. Other

References

# Sourcing Notes

The research scientists write a short set of constraints or notes about each study that are designed to help match the replication/reproduction to a team that will perform it. They are maintained as part of the project record.

No sourcing notes provided.

# Preregistration for SCORE

## Study Information

### 1. Title

Replication of a research claim from Weidmann & Callen (2013) in *British Journal of Political Science*.

### 2. Authors and affiliations

Eric L. Sevigny, PhD[1]
Jared Greathouse, BA[1]

[1] Georgia State University

### 3. Description of study

The claim selected for replication from Weidmann & Callen (2013) is that the relationship between (in)security and election fraud should be in the form of an inverted U-shape. Fraud increases with violence up to a certain level, but then decreases again (Hypothesis 1). This reflects the following statement from the paper's abstract: "It predicts (i) that the relationship between violence and fraud follows an inverted U-shape and (ii) that loyalty networks of both incumbent and challenger react differently to the security situation on the ground. Disaggregated violence and election results data from the 2009 Afghanistan presidential election provide empirical results consistent with this theory." The claim is tested with the logit models presented in Table 2. The dependent variable is a binary measure of fraud, calculated by grouping polling stations by district and applying the Beber–Scacco last-digit test to the total vote count. The dependent variable takes the value of 1 if this test is significant at the 5 percent level for a particular district. The authors measure violence as the number of attacks per capita against the International Security Assistance Force (ISAF). In Model 1 (the model selected as evidence for the SCORE program), it is included in both linear and squared form. In conjunction with a positive and significant coefficient for the linear term, the claim is tested with the coefficient on the squared term (see Model 1 for specification details). Model 1 confirms Hypothesis 1, since the linear term of the violence measure receives a positive and significant coefficient, and the squared term a negative one (coefficient on Violence (election, squared) term = -13.748, SE clustered at the regional command level = 4.720, $p < 0.01$).

## 4. Hypotheses

H1: The linear association between violence and election fraud will be positive.
H* (SCORE focal test): The quadratic association between violence and election fraud will be negative.

The SCORE focal test references the quadratic term in the original model (Table 2, Model 1). The alternative hypothesis (H1) references the linear term in the same model, which is consistent with Hypothesis 1 as specified by Weidmann and Callen (2013): “The relationship between (in)security and election fraud should be in the form of an inverted U-shape. Fraud increases with violence up to a certain level, but then decreases again.” Testing the combination of H1 and H* will more fully assess this broader claim posed in the original study.

# Design Plan

## 5. Study type

Observational Study - Data is collected from study subjects that are not randomly assigned to a treatment. This includes surveys, natural experiments, and regression discontinuity designs.

## 6. Blinding

No blinding is involved in this study.

## 7. Blinding, additional details

N/A

## 8. Study design

This replication study will employ a cross-sectional observational research design to examine the association between insurgent violence and election fraud. Consistent with Weidmann and Callen (2013), this design explores the general pattern of fraud in response to violence, which cannot produce a strong causal test of the predicted association due to the lack of exogenous variation.

The study setting remains Afghanistan as in the original study, but the current study will analyze returns from the 2014 Presidential election to replicate Weidmann and Callen's (2013) analysis of the 2009 Presidential election. It should be noted that since no candidate won a majority of votes in the April 5, 2014 election, a runoff election was held on June 14, 2014 between the top two vote-getters (Abdullah Abdullah, Ashraf Ghani). We focus on the pooled voting results for both elections in this replication to increase power for the study. The unit of analysis remains the district, which is the second-level administrative unit in Afghanistan below the province.

### 9. Randomization

N/A

## Sampling Plan

### 10. Existing data

Registration will occur prior to analysis of the data

### 11. Explanation of existing data

This study will collect, aggregate, and merge secondary data from various sources to construct a replication analysis data file. With respect to the original study, all data used for the replication reflects the updated time frame (i.e., 2014 vs. 2009).

### 12. Data collection procedures

Secondary data collection procedures for the replication analysis will be consistent with the original study. In particular, secondary data will be obtained from the following sources. Data for the 2014 Afghanistan presidential election have been made publicly available here by the National Democratic Institute (NDI), a nonprofit, nonpartisan, nongovernmental organization based in Washington, DC. Though the replication data purveyor (National Democratic Institute) is different from the original study, the source of the data is the same (International Election Commission of Afghanistan). NDI data are sourced for this project because they are provided in a more accessible digitized form. The election data provide information on election returns, polling centers, province and district names, and geographic coordinates of districts. Data on violent insurgent activity in Afghanistan have been digitized from declassified 'significant activity' reports released by US Central Command and made publicly available here by researcher Vincent Bauer (SigActs). Again, although sourced differently from the original study, the data are similarly derived from military significant activity reports. Development data from the 2013-2014 Afghanistan Living Conditions Survey (ALCS) can be requested here from the Afghanistan

government's National Statistics and Information Authority (NSIA). The ALCS is the successor to the 2007 National Risk and Vulnerability Assessment (NRVA) survey used in the original study. Elevation data (mean district elevation) comes from the World Bank here, which was derived from enhanced land elevation data produced by NASA's Shuttle Radar Topography Mission (SRTM). The original study used elevation data aggregated to the district level based on the US Geological Survey. Because elevation for nine districts was not reported in the World Bank data, we manually searched for elevation information by district latitude and longitude coordinates using Google Earth. Thus, for these nine districts, elevation is measured at one point on the map rather than as a district average from multiple locations. Distance to Kabul from each district is measured as the great circle distance on a spherical earth, identical to the original study. This distance measure was calculated within Stata using the -geodist- command based on latitude and longitude coordinates. Finally, population data have been formatted from official Afghan sources and made publicly available by researcher Colin Cookman here. Population data in the original study were based on the LandScan Global Population database.

## 13. Sample size

The unit of analysis is the district. For the 2009 Afghanistan election, Weidmann and Callen (2013) analyzed data from 398 districts (with a corresponding analytic sample of N=375 due to missing data). Various official district maps have been issued since then, some with as many as 422 districts (see here for more background information). Based on these updated district maps and election reporting results, the current replication is based on a 408 district set. However, election results were returned from just 399 districts for the first election and 401 districts for the runoff election. Thus, because we are pooling first and runoff election results for this replication, the maximum analytic sample is *N*=800. Since the unit of analysis in this replication is finite and already maximized, there will be no planned second-stage supplemental data collection and analysis with a larger sample.

## 14. Sample size rationale

Sample size is predetermined by the study design and context.

Power calculations were done in accordance with the guidelines of the Social Sciences Replication Project (SSRP). The first round of data collection aims to achieve 90% power to detect 75% of the original effect size. A second-round sample, although not planned as part of this replication, would achieve 90% power to detect 50% of the original effect size.

The power calculation for this replication can be found here: https://osf.io/n9bhj/?view_only=4d9fa88a0e504261a55875f329cdac1b.

## 15. Stopping rule

The target sample size is *N*=800, which reflects actual district-level election returns across the first (*N*=399) and runoff (*N*=401) elections. This number more than doubles the original study's maximum sample size of *N*=398 because of an increase in the number of districts and pooling of 2014 presidential initial and runoff election results. There is no plan for a second stage of data collection, and thus there is no stopping rule.

# Variables

## 16. Manipulated variables

There are no manipulated variables in this replication study.

## 17. Measured variables

Data cleaning and merging operations are provided in a Stata do file (Weidmann_Data_Cleaning.do) on the OSF project page. The final analytic dataset (Afghanistan_Election_Violence_2014.dta) is also posted on the OSF project page.

*Dependent Variable: Election Fraud*
Election data are reported at the polling station level (i.e., ballot collection points), which are nested within polling centers (e.g., mosque, school) within districts. As with the original study, voting data at the polling station level will be used to develop district-level measures of *election fraud*. We will use Beber and Scacco's (2012) 'last-digit measure' methodology to create a district-level binary dependent variable, where evidence of election fraud = '1' and lack of such evidence = '0.' In particular, election fraud is indicated when the within-district distribution of the last digit among polling station vote totals deviates significantly ($p<0.05$) from the expected uniform distribution. The use of Beber and Scacco's (2012) last-digit method exactly replicates the operationalization in the original study.

*Focal Independent Variable: Violence*
Weidmann and Callan (2013) measured district-level insurgent violence as "the number of attacks per capita against the International Security Assistance Force (ISAF), [defined] as the total of insurgent-initiated direct fire incidents, incidents involving improvised explosive devices (IEDs), and small arms fire [occurring] in a five-day window around election day." Like Weidmann and Callen (2013), replication data on *insurgent violence* in Afghanistan come from declassified military 'significant activity' or SigAct reports released by US Central Command. The data used for the replication have been collected, digitized, and made publicly available by researcher Vincent Bauer. These data comprise *N*=431,547 geographically and temporally referenced events

for the period January 2008 to December 2014. The replication study will follow Weidmann and Callen's (2013) operationalization of insurgent violence, defined as the number of insurgent violent events occurring in a five-day window surrounding election day (i.e., April 5-9, 2014 for first election and June 14-18, 2014 for runoff election) at the district level. Similar to the original study, insurgent violent events are defined as enemy actions involving (i) assassinations, direct fire, or small arms fire, or (ii) IED/mine explosive hazards.

In addition to the main results reported in Table 2, Model 1, Weidmann and Callen (2013) also reported in Table 2, Model 2 a supplemental analysis measuring violence during a wider sixty-day window prior to election day. We will replicate this supplemental analysis using the extended window of observation. Consistent with Weidmann and Callen (2013), both violence measures will be aggregated to the district level and operationalized as incidents per 1,000 population.

*Control Variables*

*Percentage of planned polling centers closed*. This measure is operationalized from NDI election data as the number of closed polling centers on election day divided by the total number of planned polling stations prior to election day, multiplied by 100. This measure differs from Weidmann and Callen (2013), who measured the *number of closed polling stations*. We deviate from this measure for two reasons. First, data on the number of planned polling stations is unreliably measured; we have more confidence in data on planned polling centers. Second, normalizing by the number of planned centers adjusts for underlying district population differences as reflected by number of polling centers. Weidmann and Callen (2013) did not normalize polling station counts because each station is limited to collecting 600 ballots, so normalization was inherent in the measure. This is not the case with polling centers, which can support variable numbers of polling stations. We therefore normalize the measure as a percentage to make the measure meaningful in context.

*Electrification*. Weidman and Callen (2013) developed a measure of *proportion of households with electricity* from the 2007 National Risk and Vulnerability Assessment (NRVA) survey, administered by the Ministry for Rural Rehabilitation and Development (MRRD) and the Central Statistics Office (CSO) of Afghanistan, and based on information from *N*=20,576 households in 383 districts. A similar measure of electrification will be created from the 2013-2014 Afghanistan Living Conditions Survey (ALCS), the successor to the NVRA, which was based on information from *N*=20,786 households in 364 districts. Although Weidmann and Callen (2013) do not explicitly describe how they operationalized this measure from NRVA questionnaire items, we will operationalize this measure as the survey weighted percentage of households within each district answering affirmatively to any response option for the following question: "Has your household had electricity at any time in the past month from any of these sources?: Electric grid, Government generator, Private generator (engine), Private generator (hydro), Community generator (engine), Community generator (hydro), Solar, Wind, Battery." Our operationalization differs from Weidman and Callen (2013) in scaling: rather than a proportion (0-1) we express this measure as a percentage (0-100) to improve model interpretation.

*Per capita expenditure*. Weidmann and Callen (2013) also developed a measure of *per capita expenditure* by households using the NRVA, but again were not explicit about how they operationalized this measure from NRVA questionnaire items. We will operationalize a weighted measure of past-month *per capita expenditures* (in Afghanis) using the 2013-14 ALCS. In particular, we will operationalize this measure as follows:

1. Sum the reported past-month household expenditures ("How much did this household spend in the last month for each type of fuel used in the household?: Electricity; Gas; Fuel oil; Firewood; Ping, straw, manure") and other household costs ("In the past month, how much did your household spend on the following categories of items? Food at home; Food away from home; Grooming, cosmetics, laundry soap, shampoo; Housing rent; Communication (phone, internet); Transportation of household members").
2. Sum the reported past-year household expenditures for other household expenses ("In the past year, how much did your household spend on the following categories of items?: Clothing and shoes; Health services and medicines; Education (tuition, books, uniform, supplies"), and then divide by 12 to obtain average monthly expenditures.
3. Sum the past-month and average past-month household expenditure amounts from steps 1 and 2.
4. Sum the survey weighted total expenditure amounts by district.
5. Divide by 1,000 to express expenditures in thousands of Afghanis.
6. Divide by the district population to express as per capita household expenditures.

*Distance to Kabul*. Data on the distance (km) from each district to Kabul is measured as the great circle distance on a spherical earth, as with the original study. This measure will be operationalized as the straight-line distance from each district center to the capital of Kabul based on latitude and longitude coordinates. In particular, we used the user-written Stata command -geodist- to calculate these distances.

*Elevation*. *Average elevation* data (in kilometers) for each district are based on World Bank staff estimates of 2015 satellite data from NASA's Shuttle Radar Topography Mission (SRTM), supplemented for missing data in nine districts by using Google Earth queries by geographic coordinates. By comparison, Weidmann and Callen (2013) calculated elevation from US Geological Survey data. Our operationalization differs from Weidman and Callen (2013) in scaling: we express as kilometers rather than meters to improve model interpretation.

*Population*. Population data used to normalize other measures are captured at the district level to a specificity in the thousands based on official government statistics.

*Clustering variable*. Weidmann and Callen (2013) clustered their analyses by coalition forces regional commands (RCs). In 2009, the 34 provinces were clustered into one of four RCs. By 2014, the number of RCs had expanded to six (see here). Our main clustering variable keys on these six RCs.

## 18. Indices

N/A

# Analysis Plan

## 19. Statistical models

For the purposes of SCORE, to test the combination of **H1** and **H*** in line with the original paper (Table 2, Model 1), we will use logit models to regress the last-digit fraud measure on violence and its squared term to test the inverted U-shaped hypotheses, while controlling for covariates indicated above. The measure of violence in this model is the incidence rate per 1,000 population within a 5-day window around election day.

We will also perform a supplemental analysis using the alternative measure of violence defined as the incidence rate per 1,000 population during a 60-day window preceding election day (Table 2, Model 2).

For both analyses, we will estimate cluster-robust standard errors by regional command, consistent with Weidmann and Callen (2013). We will also estimate supplemental models using multiway clustering grouped by regional command and election cycle to additionally account for pooling of election results.

All analyses will be performed using Stata MP 16.1. Analytic code is provided on the OSF project page in a Stata do file (Weidmann_Data_Analysis.do). Test results of the planned analyses on a 5% random sample of the data is provided on the OSF project page. The results demonstrate the analytic code works properly.

## 20. Transformations

We will use a quadratic (i.e., squared) term for the focal independent variable, violence, to test the nonlinear hypothesis that election fraud will increase as violence increases to a point, whereupon election fraud will decline at higher levels of violence.

## 21. Inference criteria

Criteria for a successful replication attempt for the SCORE project is a statistically significant effect ($p < .05$, two tailed) in the same pattern as the original study on the focal hypothesis tests (**H1** and **H***). For this study, this criteria is met by a positive and significant coefficient on the linear

term for election violence (H1) and a negative and significant coefficient on the squared term for election violence (H*) in the focal regression model.

## 22. Data exclusion

We do not have any *a priori* exclusion criteria for observations in our sample.

## 23. Missing data

Consistent with Weidmann and Callen (2013), we will perform a complete case analysis. In the original study, this resulted in the loss of 6% of the data (375/398). In the current replication, a complete case analysis is expected to result in the loss of 76 observations, or less than 10% of the data (724/800), due to missingness on the ALCS measures.

## 24. Exploratory analysis

N/A

## 25. Other

The final aggregated data set will be made publically available on the OSF project page.

## References

Bradley, Walker and Brian Blankespoor. (2019). Measuring Mobility in Afghanistan Using Travel Time Models Methodology Note. Washington, D.C.: World Bank Group.

Weidmann, Nils B. and Michael Callen. (2013). Violence and Election Fraud: Evidence from Afghanistan. *British Journal of Political Science*, 43, 53-75.