Make Two Democracies and Call Me in the Morning: Endogenous Regime Type and the Democratic Peace *

Daina Chiba
University of Essex
dchiba@essex.ac.uk

Erik Gartzke
University of California, San Diego
egartzke@ucsd.edu

May 6, 2020

Abstract

The rarity of violent conflict between pairs of democracies has often been attributed to a pacifying effect of democracy. As democracy and peace may have common causes, however, scholars have also cautioned against interpreting the observed correlation as a causal relationship. Estimating the causal effect of democracy on peace using observational data has been challenging because some of the common causes of the two may be unobservable. We propose an approach that addresses the endogeneity issue using instrumental variables. Using average fertility rate as an instrument, we find that democracy does not have a pacifying effect. If anything, democratic countries are more likely to attack other countries. Our findings call for a serious reconsideration of the theories that have been proposed to explain the observed peace between democracies.

Word count: 11,888 words (without appendix)

*Paper prepared for delivery at Asian Online Political Science Seminar Series to be held on 20 May, 2020. We thank Nils-Christian Bormann, Kevin A. Clarke, Shohei Doi, Songying Fang, Kristian Skrede Gleditsch, Håvard Hegre, Jesse C. Johnson, Kyosuke Kikuta, Atsushi Tago, Yuki Yanai, and Sebastian Ziaja for valuable feedback and suggestions.

The democratic peace — an observation that pairs of democracies experience militarized conflict with much lower frequency than other pairs do — has been one of the most prominent empirical regularities, stimulating a lot of scholarly and policy debates for the past decades. Although a general consensus has emerged among scholars as to the correlation between joint democracy and peace, a considerable disagreement remains as to the cause(s) of this empirical regularity. Numerous theories have been proposed to explain why democracy may cause peace, or alternatively, why peace may cause democracy (= reverse causality), or why some third factor may cause both democracy and peace, making it look that these two are related (= spurious correlation).

We seek to contribute to this debate not by proposing yet another theory to explain the democratic peace but by estimating the causal effect of joint democracy on the probability of militarized conflict using a quasi-experimental research design. In so doing, we acknowledge that some of the common causes of democracy and peace may be unobservable, generating endogeneity in the relationship between the two. We correct for this endogeneity via an instrumental variables framework. Our findings suggest that the effect of democracy on conflict is either positive or indistinguishable from zero depending on specifications and the baselines. The results call for a serious reevaluation of the theories proposed to explain why democracy reduces conflict.

The rest of the article proceeds as follows. In the next section we briefly review previous attempts to estimate the effect of democracy on conflict and point out methodological issues that have plagued the previous efforts. We then introduce our empirical strategy and validate it via Monte Carlo simulations. The following section discusses the main findings and a series of falsification tests to assess the plausibility of assumptions necessary to facilitate a causal inference with observational data. The final section concludes with recommendations for future research.

The Democratic Peace Debate

Although few scholars dispute the existence of the correlation between joint democracy and peace, a vigorous debate arises over whether democracy is the cause of this correlation. We can identify three competing clusters of arguments proposed to explain this empirical regularity. The first cluster of explanations posit that democracy causes peace. Scholars have identified various aspects of democracy as the cause of peace, including shared norms of peaceful dispute resolution (Maoz and Russett 1993), democratic institutions that constrain leader's war-making abilities (Bueno de Mesquita et al. 1999), and more credible communication thanks to greater transparency (Schultz 1998; Choi and James 2007; Colaresi 2014) or greater audience costs (Fearon 1994; Schultz 1999).

The second cluster of explanations posit that the causal arrow goes the other way around — peace causes democracy. Scholars argue that democratization is more likely to occur in a more peaceful international environment, because a threatening external environment requires states to allocate their resources to defence and centralization of powers. Within this cluster, some scholars partially agree with the first cluster argument, proposing that the causal arrow runs both ways — peace causes democracy but democracy causes peace as well (Gleditsch 2002b; Rasler and Thompson 2005; Reuveny and Li 2003; Thompson 1996). Others are more critical of the first cluster, claiming that peace causes democracy but not the reverse (James, Solberg, and Wolfson 1999; Gibler and Tir 2014).

Finally, the third cluster of explanations posit that a confounding variable causes both democracy and peace, thus creating a spurious correlation. Scholars have identified various factors such as Cold War alliances (Farber and Gowa 1997; Gowa 1999), common national interests (Gartzke 1998, 2000), capitalist economy (Gartzke 2007; Gartzke and Hewitt 2010),[1] contractualist economy (Mousseau 2013, 2018),[2] stable territorial borders (Gibler

[1] See also Choi (2011) and Dafoe (2011).
[2] See also Dafoe, Oneal, and Russett (2013).

(a) Confounding Z (b) Post-treatment Z

Figure 1: This figure illustrates two commonly presumed scenarios involving democracy, peace, and a third variable, Z. In scenario (a), Z influences both democracy and peace. In scenario (b), democracy influences Z, which in turn influences peace.

2007, 2012),[3] and many others as a possible confounding variable that could explain away the democratic peace.

Each of these three clusters of explanations claims to find empirical support for their argument. Given that these studies have typically relied on similar (or often identical) sets of observational data, one might wonder why different scholars have drawn different conclusions. In our view, there are two methodological dilemmas at the heart of observational studies on the democratic peace that have allowed us to draw contradictory inferences from the same data.

The first dilemma arises from the fact that scholars often disagree over the role of third variables in the democracy–peace nexus. Consider a particular determinant of peace, Z. If one believes that Z also causes democracy, then Z is a potential confounding variable, corresponding to the diagram (a) on the lefthand side of Figure 1. As such, we must control for Z in assessing the relationship between democracy and peace because a failure to control for Z would result in an endogeneity (omitted variable) bias. If, on the other hand, one believes Z is caused by democracy, then Z is a post-treatment variable, corresponding to the diagram (b) on the righthand side of Figure 1. As such, we must *not* control for Z in assessing the relationship between democracy and peace because mistakenly controlling for Z would result in a post-treatment variable bias. Depending on the direction of the causal arrow between democracy and Z, the correct treatment of Z will thus be different. This poses a dilemma – if we control for a particular Z, we reduce the risk of endogeneity

[3] See also Park and Colaresi (2014) and Gibler (2014).

bias, but this comes with the cost of increased risk of post-treatment variable bias; if we omit this Z, on the other hand, we reduce the risk of post-treatment variable bias, but this comes with the cost of increased risk of endogeneity bias.

Suppose that we find evidence for a statistical association between democracy and peace in a bivariate analysis, but that this statistical association disappears when we control for a particular third variable, Z. If we believe that this Z is a confounding variable, we must conclude that the democracy–peace relationship is spurious. If, however, we believe that this Z is a post-treatment variable, we can still conclude that democracy causes peace through Z. Consider the scholarly debate over the role of states' preference similarity in assessing the democratic peace. One of the major disagreements between the two sides of the debate is about whether preference similarity is a confounding variable (preference similarity leads to regime similarity) (Gartzke 1998, 2000) or a post-treatment variable (regime similarity leads to preference similarity) (Oneal and Russett 1999). If one subscribes to the former view, the observation that the statistical association between democracy and peace disappears when controlling for state preference similarity should lead one to conclude that this is a spurious correlation. If one subscribes to the latter view, however, the very same observation allows one to conclude that democracy causes peace.

The second dilemma is that, no matter how many potential confounding variables we control for, there is always a risk that some unobserved or unobservable determinant of democracy and peace is still omitted from the analysis, generating endogeneity bias. Scholars have employed fixed-effects models to control for unit-specific (Green, Kim, and Yoon 2001; Schultz 2001), time-invariant unobserved confounders, but such methods will not eliminate endogeneity caused by unobserved confounders that vary over time. Fixed-effects models also tend to mask the impact of slow-changing variables such as regime type (Beck and Katz 2001).

Looking for Causality

In face of these dilemmas, experimental approaches — randomized controlled trials, in particular — are the most powerful methods to estimate causality. Randomization of treatment assignment, if possible, would allow us to eliminate any bias that may arise from unobservable and observable confounding variables. That being said, the limitation of experimental methods is that treatment randomization is not always possible due to feasibility and ethics considerations. For example, an analyst cannot take actual countries as experimental subject, randomly assign regime type to them, and observe war and peace as the outcome variable.

Experimental studies on the democratic peace have thus focused on testing some components of the theories by disaggregating the causal chain between regime type and peace into smaller pieces. In particular, scholars have conducted survey experiments to explore the role of public opinion in the causal mechanisms (Mintz and Geva 1993; Johns and Davies 2012; Tomz and Weeks 2013). These studies seek to identify the causal effect of regime type on public support for military actions, which could, in turn, influence war and peace. By randomly assigning the regime type of the target country in a hypothetical military action in different experimental scenarios, they find that democratic publics tend to oppose military actions against democracies than against non-democracies.[4]

Although this finding has the potential to be a part of the explanation of the democratic peace, it alone is insufficient. Replicating Tomz and Weeks's (2013) survey experiment in a non-democracy (China), Bell and Quek (2018) find that non-democratic publics also tend to oppose military actions against democracies than against non-democracies. If both non-democratic and democratic publics oppose fighting with democracies, democracies should have lower probability of conflict not only with democracies but also with non-democracies. Therefore, the difference in public opinion cannot be the cause of the

[4] Tomz and Weeks (2013) taps further into the causal mechanisms by exploring *why* democratic publics tend to oppose military actions against democracies more than they do against non-democracies.

difference in conflict propensity between democracy-democracy pairs and other pairs. [5] Moreover, the relationship between preference and observed outcome could be more complicated due to strategic interaction. For example, if countries in general prefer to avoid fighting with democracies, this may embolden democracies in crisis bargaining, which may in turn lead to an increase in probability of conflict for pairs involving democracies.

One promising way forward to overcome the shortcomings of the observational studies and experimental studies would be to combine the strengths of both approaches. Specifically, we propose to use a quasi-experimental design — in particular, instrumental variables (IV) framework — to identify the causal relationship in the democratic peace, using the commonly used observational data. Such an approach allows us to study the overall relationship between regime type and peace (rather than intermediate outcomes such as levels of public support) and to identify the causal effect of one on the other in a credible manner.

Before we detail our approach, it is worth while to discuss why such an attempt has not been done successfully in the past and what makes ours different from past work. We are not the first to apply an IV framework (more specifically) or multi-equation models (more broadly) to the democratic peace. However, previous attempts suffer from two major problems. First, previous studies have typically used a dyad (country pair) as the unit of observation in analyzing conflict, which requires some summary measure(s) of democracy for a pair of countries rather than the original democracy measure for each country. [6] This creates a discrepancy between the first stage regression (predicting democracy at the country level) and the outcome stage regression (predicting conflict at the dyad level). [7]

[5] It could be the government's responsiveness to public opinion, rather than public opinion itself, that explains this difference. However, this is an entirely different causal mechanism than what Tomz and Weeks (2013) propose.

[6] The most commonly implemented way to summarize democracy values for the two countries in a dyad is to take the smaller of the two values that measure the degree of democracy, often coupled with the larger of the two values. Some studies alternatively use the product (i.e., an interaction term) of two democracy scores, especially when using dichotomous measures of democracy.

[7] For example, Reuveny and Li (2003) and McDonald (2015) both use dyad (-year) as the unit of observation in their multi-equation models of the democratic peace, with dyadic conflict as the outcome variable and the lower democracy as an endogenous independent variable (Reuveny and Li (2003) use

We avoid this problem by using *directed* dyad as the unit of observation in predicting conflict, distinguishing between the potential challenger and target in a dispute. This allows us to connect the first stage equations (predicting challenger's and target's regime types) and the outcome stage equation seamlessly. Doing so has several benefits: the outcome stage model could directly include country-level covariates (such as challenger's and target's democracy) without forcefully converting them to a dyadic summary; we could also estimate the system of equations jointly rather than relying on the forbidden regression.

Second, a more daunting challenge in applying an IV approach to the democratic peace research is the difficulty of finding a plausible instrument for regime type — a variable that is strongly correlated with regime type but is unrelated to the outcome variable of interest, i.e., conflict. This is the challenge that has plagued empirical researchers in many fields. For example, a recent study on the effect of regime type on economic growth uses a diffusion-based measure of democracy (i.e., average value of democracies in a given region) as an instrument for democracy (Acemoglu et al. 2019). However, diffusion-based instruments such as this are unlikely to be a valid instrument, due to spatial spill-over, interdependence, and, most importantly, simultaneity (Betz, Cook, and Hollenbach 2018). Recognizing problems of spatial instruments, McDonald (2015) seeks to exploit the very discrepancy between country-level and dyad-level designs as the source of identification. His discussion, however, lacks a clear explanation as to why some determinants of regime type do not influence conflict. [8]

---

the higher democracy as well). The first stage regression uses country-year as the unit of observation and predicts democracy scores. They then calculate the lower (and the higher) of the predicted values of democracy scores for each dyad-year and use them as predictors in the outcome stage probit regression predicting dyadic conflict. Such an approach has at least two problems. First, this two-step method is an example of "forbidden regression" (Angrist and Pischke 2009) — plugging in predicted values is forbidden unless the outcome stage is a linear model. Second, it is unclear whether and how the first-stage predictors (covariates of country-level democracy) should be included in the dyadic-level outcome stage model, and (if included) how to interpret them.

[8] McDonald's (2015) IV model (presented in Table 6 of the online Supplementary Appendix) requires an instrument that strongly influences regime type but is unrelated to conflict. However, he claims that his instruments (e.g., great power alliance of one side of a dyad) are a strong determinant of conflict but are unrelated to the regime type of the other side in a dyad (pages 4–5 of Supplementary Appendix). In other words, he claims his instruments have no relevance (uncorrelated with the endogenous covariate) and do not have exogeneity (correlated with the outcome), whereas he would need to claim

We turn to a demographic variable — average female fertility rate in a given country — as a source of variation in regime type that is exogenous to international conflict. As we will argue below, a declining fertility rate is a strong driver of democratization. We will also present theoretical arguments and a series of falsification tests that support the claim that fertility rate does not directly influence international conflict.

Research Design

We will first present our statistical model that allows us to estimate three discrete-choice equations jointly. In doing so, we will illustrate how a separate estimation could generate endogeneity bias due to omitted variables. We will then discuss our data and identification strategy.

The Model

Let y_i denote conflict initiation by a potential challenger country c_i against a potential target country t_i in a directed dyad (challenger-target pair) i, where the value of 1 is observed when conflict is initiated and 0 otherwise. [9] Suppose we have binary measures of regime type for the challenger and target, D_i^c and D_i^t, coded 1 for democracies and 0 for non-democracies. A conventional approach to obtain the effect of joint democracy on conflict is to estimate the following regression:

$$y_i = 1[\beta_1 D_i^c + \beta_2 D_i^t + \beta_3 D_i^c D_i^t + \boldsymbol{x_i}\boldsymbol{\beta_x} + \eta_i > 0], \tag{1}$$

where β_1, β_2, and β_3 are coefficients for democracy measures, $\boldsymbol{x_i}$ is a vector of exogenous covariates for i, $\boldsymbol{\beta_x}$ is a vector of coefficients for $\boldsymbol{x_i}$, and η_i is an error term. Then, the

the opposite to justify his instruments.

[9] Throughout the presentation of the model and Monte Carlo simulations, we will omit the time component of observations for simplicity.

estimated average causal effect of joint democracy for a sample of size n, $\hat{\tau}$, is obtained as:

$$\hat{\tau} = \sum_{i=1}^{n} \{F_\eta(\hat{\beta}_1 + \hat{\beta}_2 + \hat{\beta}_3 + \boldsymbol{x_i}\hat{\boldsymbol{\beta}}_{\boldsymbol{x}}) - F_\eta(\boldsymbol{x_i}\hat{\boldsymbol{\beta}}_{\boldsymbol{x}})\}, \tag{2}$$

where $F_\eta(\cdot)$ is the cumulative distribution function for the error term.[10] If joint democracy is a cause of peace, $\hat{\tau}$ will be negative, which is equivalent to $\hat{\beta}_1 + \hat{\beta}_2 + \hat{\beta}_3 < 0$. [11] However, the causal effect estimated from equation (1) this way may suffer from endogeneity bias if there are unmeasured factors influencing both conflict (y_i) and regime type (D_i^c and D_i^t). For example, when some unmeasured factors encourage democracy and discourage conflict, D_i^c and D_i^t tend to take a higher value in cases where η_i takes a lower value. Lower values of the error term will then be absorbed by the coefficients for democracies. This will thus generate a downward bias to the estimates of β_1, β_2, and β_3, making it more likely to find evidence in support of pacifying effect of democracy even when it does not exist in truth.

To correct for this endogeneity problem, we model y_i, D_i^c, and D_i^t as a joint process using a trivariate model. The model has three equations connected via the trivariate Gaussian copula, thus allowing us to capture the correlation between the error terms for the equations. We have the following system of equations:

$$y_i = 1[\beta_1 D_i^c + \beta_2 D_i^t + \beta_3 D_i^c D_i^t + \boldsymbol{x_i}\boldsymbol{\beta}_{\boldsymbol{x}} + \eta_i > 0] \tag{3}$$

$$D_i^c = 1[\boldsymbol{z_i^c}\boldsymbol{\gamma_c} + \epsilon_i^c > 0] \tag{4}$$

$$D_i^t = 1[\boldsymbol{z_i^t}\boldsymbol{\gamma_t} + \epsilon_i^t > 0], \tag{5}$$

[10] As the outcome variable, conflict initiation, is a rare event, we use the complementary log-log (clog-log) specification for the conflict equation throughout this study.

[11] A commonly accepted variant of the democratic peace argument further stipulates that democracies are less conflict-prone only toward other democracies but that democracies are in general not less conflict-prone toward non-democracies. This suggests that $\hat{\beta}_1$ and/or $\hat{\beta}_2$ may be positive or zero but $\hat{\beta}_3$ is negative and greater in absolute value than $\hat{\beta}_1$ and $\hat{\beta}_2$ combined. In other words, this argument implies $\hat{\beta}_3 < -(\hat{\beta}_1 + \hat{\beta}_2) < 0$. The following discussion equally applies to this variant of the democratic peace.

where $\boldsymbol{z_i^c}$ and $\boldsymbol{z_{ci}^t}$ are vectors of covariates for challenger's and target's regime type in dyad i, respectively, $\boldsymbol{\gamma_c}$ and $\boldsymbol{\gamma_t}$ are vectors of coefficients for variables in $\boldsymbol{z_i^c}$ and $\boldsymbol{z_i^t}$. Let $\epsilon_i^y = \Phi^{-1}\{F_\eta(\eta_i)\}$, where $\Phi^{-1}(\cdot)$ denotes the quantile function for the standard Normal distribution. Then, ϵ_i^y and the error terms for the second and third equations follow the trivariate Normal distribution $(\epsilon_i^y, \epsilon_i^c, \epsilon_i^t)' \sim^{iid} \mathcal{N}_3(0, \Sigma)$, where [12]

$$\Sigma = \begin{bmatrix} 1 & \rho_{12} & \rho_{13} \\ \rho_{21} & 1 & \rho_{23} \\ \rho_{31} & \rho_{32} & 1 \end{bmatrix}.$$

On the diagonal are the error variance normalized to 1, and the off-diagonal elements are the correlation between the error terms such that $\rho_{jk} = \rho_{kj}$ for $j \neq k$.

Our approach to model democracies and conflict jointly is an application of instrumental variables (IV) regression. [13] Equations (4) and (5) above jointly constitute the "first stage" of an IV estimation, where endogenous predictors, D_i^c and D_i^t, are regressed on covariates. Equation (3) is the "second stage" or the outcome stage, where the outcome of interest, y_i, is regressed on endogenous predictors and other covariates.

We conduct Monte Carlo simulations to illustrate how a univariate estimation of equation (1) produces bias in the estimate of the causal effect of democracy on conflict and that a trivariate estimation of equations (3)–(5) recovers the correct causal effects. We do this exercise because using the trivariate model as an IV model is not commonly done in political science. [14] The results of the simulations are presented in the Appendix.

[12] In other words, our model is a trivariate model where the first equation has the clog-log link and the two other equations have the probit link.

[13] As such, the causal effect we identify is a *local* average treatment effect (LATE), rather than an average treatment effect (ATE). Although LATE is not necessarily equal to ATE, this is the best we can hope for when we cannot perform treatment randomization. We will revisit this later in the paper.

[14] We are aware of one application in the field of economics (Filippou et al. 2019) where the trivariate probit model is used to endogenize two binary predictors.

The Data

We analyze observational data on democracy and conflict for the period between 1960 and 2010. [15] Our sample includes all contiguous directed dyad years identified by the Correlates of War (COW) data. Directed dyads are country pairs comprised of a potential challenger and a potential target. Distinguishing between the challenger and the target allows us to study how the regime type of each side influences the risks of militarized conflict initiation and targeting. We focus on contiguous country pairs (i.e., countries either share a land border or are separated by less than 400 miles of water [16]) to reduce the heterogeneity in the sample and make each observation unit more comparable. [17] The vast majority of non-contiguous country pairs do not interact with each other on a regular basis and thus have no opportunities to experience militarized conflict. Excluding such dyads from our analysis thus enhance the internal validity of our study, although it may come at a cost of limiting the generalizability of our findings beyond contiguous dyads. [18]

We identify whether the potential challenger in a dyad initiates a militarized conflict against the target in a given year, based on COW's Militarized Interstate Dispute (MID) dataset. We use a new version of MID dataset compiled by Gibler, Miller, and Little (2016) that introduces a number of corrections to version 4.01 of MID. [19] To reduce heterogeneity in the MID data, we omit "protest-dependent" MIDs that are based on a substantially different coding logic (Gibler and Little 2017). [20] Our outcome variable, conflict initiation,

[15] We choose this observation period because of the data availability of important variables. Data on fertility rate are available only since 1960, and those on conflict are available only up to 2010.

[16] The data are based on version 3.2 of the COW Direct Contiguity data set (Stinnett et al. 2002).

[17] Reed and Chiba (2010) report evidence that suggests there is a considerable heterogeneity between contiguous dyads and non-contiguous dyads not only in terms of covariate characteristics but also in terms of how these dyads respond to a change in covariate values. More specifically, their findings suggest that the effect of regime type on conflict differs between contiguous and non-contiguous dyads.

[18] As a robustness check, we analyze "politically relevant" dyads (i.e., sample includes all contiguous pairs plus non-contiguous pairs involving at least one major power states) and the results are consistent. The results will be included in the Appendix in the next round of revisions.

[19] We use version 2.1.1 of the Gibler-Miller-Little MID data available online at `http://svmiller.com/gml-mid-data/` (last accessed in March 2020). See also Palmer et al. (2019) and Gibler, Miller, and Little (2020).

[20] An alternative approach to reduce heterogeneity of MIDs is to focus on those MIDs that involve fa-

Table 1: Two-way frequency table of democracy measures

		Polity			
		Democracy	Non-democracy	NA	Sum
BMR	Democracy	2560	306	406	3272
		(33.5)	(4.0)	(5.3)	(42.9)
	Non-democracy	148	3923	288	4359
		(1.9)	(51.4)	(3.8)	(57.1)
	Sum	2708	4229	694	7631
		(35.5)	(55.4)	(9.1)	(100.0)

Cell entries are the number of country-years. Cell-percentages are in the parentheses.

is coded as 1 for a directed-dyad-year if the challenger initiates a new MID against the potential target, 0 otherwise, and missing if an MID is already ongoing (McGrath 2015).

We use Boix, Miller, and Rosato's (2013) dichotomous coding of democracy data as our primary measure of regime type. [21] This measure, hereafter BMR, distinguishes between democracy and non-democracy by looking at contestation (free and fair elections) and participation (suffrage). Although different theories of democratic peace focus on different aspects of democracy, focusing on contestation and participation is probably one of the most uncontroversial common denominators of different definitions of democracy. BMR's dichotomous nature facilitate an intuitive interpretation of the causal effect of democracy, and BMR also has an advantage over other measures of regime type in terms of coverage. That said, to check the robustness of our findings, we also prepare an alternative measure based on the more conventional polity2 variable from the Polity IV project (Marshall, Jaggers, and Gurr 2018). [22] Following the convention, we code a country as a democracy if it receives the polity2 score of 6 or higher.

Table 1 is a cross tabulation of these two democracy measures for all country-year ob-

talities (Hegre 2000). However, such escalation-based selection is problematic, as it prevents us from understanding the causes of disputes that were de-escalated before fatalities were incurred.

[21] We use version 3 of the data, updated to cover up to 2015. The data are available at https://dataverse.harvard.edu/dataverse/BMR (Last accessed on August 6, 2019).

[22] The data are available at http://www.systemicpeace.org/polityproject.html (Last accessed on August 6, 2019).

servations in our analysis sample. [23] For about 85% of the observations, these two measures agree with each other. Disagreement occurs for about 6% of the observations, and the Polity measure is missing for the remaining 9% of the observations. Tables A.1 and A.2 in the Appendix list all cases of disagreement.

In estimating the conflict equation of the model, we control for factors that could influence the probability that a challenger initiates conflict against a target. We expect that a challenger is more likely to initiate a militarized conflict when it is more likely to prevail in battles. To capture this expectation, we include challenger's military advantage operationalized as the challenger's military capability score (COW's CINC score) divided by the sum of CINC scores for challenger and target. As economically more developed countries will have more resources to allocate to military actions, economic development may influence conflict involvement. We thus control for both challenger's and target's economic development measured by per capita GDP. [24] Finally, to control for duration dependence of military conflict, we include the natural log of the number of years since last involvement in a militarized dispute.

Identification: fertility rate as an instrument

To identify the causal effect of regime type on conflict using the IV framework, we need an instrument (instrumental variable) that is both *relevant* (i.e., strongly correlated with democracy) and *exogenous* (i.e., unrelated to conflict except through its effect that goes through democracy).In other words, $\boldsymbol{z}_i^c$ and $\boldsymbol{z}_i^t$ in equations (4) and (5) explaining democracy should include a significant determinant of democracy that does not belong in $\boldsymbol{x}_i$ in equation (3) explaining conflict. Finding an instrument that is both relevant and exogenous is extremely difficult in empirical research. Many well-known correlates of democracy, such as economic development, do not satisfy the exogeneity condition as we could

[23] This includes all countries recognized by COW, except for those that are separated from other COW countries by more than 400 miles of water (e.g., Iceland and New Zealand) and those for which other covariates are missing.

[24] The data on GDP and population are from Gleditsch (2002a).

think of plausible theoretical mechanisms linking them with conflict.

We argue that fertility rate (the average number of children that a woman would have over her lifetime) in a country serves as a valid instrumental variable in estimating the effect of democracy on conflict. We start by showing theoretical and empirical evidence that this is a relevant instrument. After presenting our argument for instrument exogeneity, we first present and discuss the main empirical results. We will then return to the issue of instrument exogeneity more thoroughly.

Instrument relevance

We point out two causal mechanisms through which lower fertility rate leads to an onset and stabilization of democratic government (Sommer 2018; Wilson and Dyson 2017). First, a decline in fertility rate will lead to economic and political empowerment of women, which in turn leads to improved gender equality in a society (Inglehart and Norris 2003; Iversen and Rosenbluth 2010). As women are freed up from the heavy burdens of childbirth and childcare, they can invest more time and resources on schooling and salaried work. Women with higher education and a stable employment will have a better access to information and resources necessary to participate in politics. With more women becoming politically active, women's descriptive and substantive political representation will improve. Gender equality in politics (i.e., better representation of female rights and better political participation of women) is an essential component of democracy.

Second, a lower fertility leads to greater education spending per capita both at the macro and micro levels, which in turn supports democracy. At the macro level, having fewer children in a society for a fixed level of government spending on education amounts to improved education spending per capita. At the micro level, declining fertility rate means that each family will have more time and resources to spend on educating their children. In other words, when it becomes no longer necessary for families to have many children to secure sufficient labor force, they will be able to allocate their resources on

education. The effect of improvement in education on political participation is well established in the literature (Inglehart and Norris 2003). Furthermore, a recent study by Paglayan (2018) shows that improvement in education is a cause, rather than a consequence, of democratization.

There is a growing body of literature that examines the causal link between demographic transition and the emergence of democracy (e.g., Dyson 2013; Wilson and Dyson 2017). In particular, conducting a systematic examination of this relationship, Sommer (2018) finds that fertility rate is strongly associated with the emergence of democracy even after controlling for other important determinants of regime change and for reverse causality. Turning to our analysis sample, there are many democracies that have low fertility rate (e.g., most of Western advanced economies) and non-democracies that have high fertility rate throughout the observation period 1960-2010.[25] As we will show later in the paper, fertility rate is indeed a powerful predictor of democracy in our sample. More importantly for our purpose of illustrating the causal link between the two, there are many countries that (i) have a high fertility rate and are non-democratic during the first part of the observation period, (ii) experience a large decline in fertility rate, and (iii) the decline in fertility rate is followed by a democratic transition. Figure 2 shows some of those countries that satisfy conditions (i)–(iii).

Instrument exogeneity

For fertility rate to be a valid instrument, it must also satisfy the exogeneity condition [26] — fertility must not have any direct effect on conflict, that is, any effect of fertility on conflict must go through regime type. Although this is an untestable assumption, we seek to assess its plausibility by contemplating on possible direct theoretical connections between

[25] Data on fertility rate are obtained from the Gender Statistics database from the World Bank, available at: https://datacatalog.worldbank.org/dataset/gender-statistics (last accessed on 30 March, 2020).

[26] The exogeneity condition is often called the exclusion restriction. We use these two terms interchangeably.

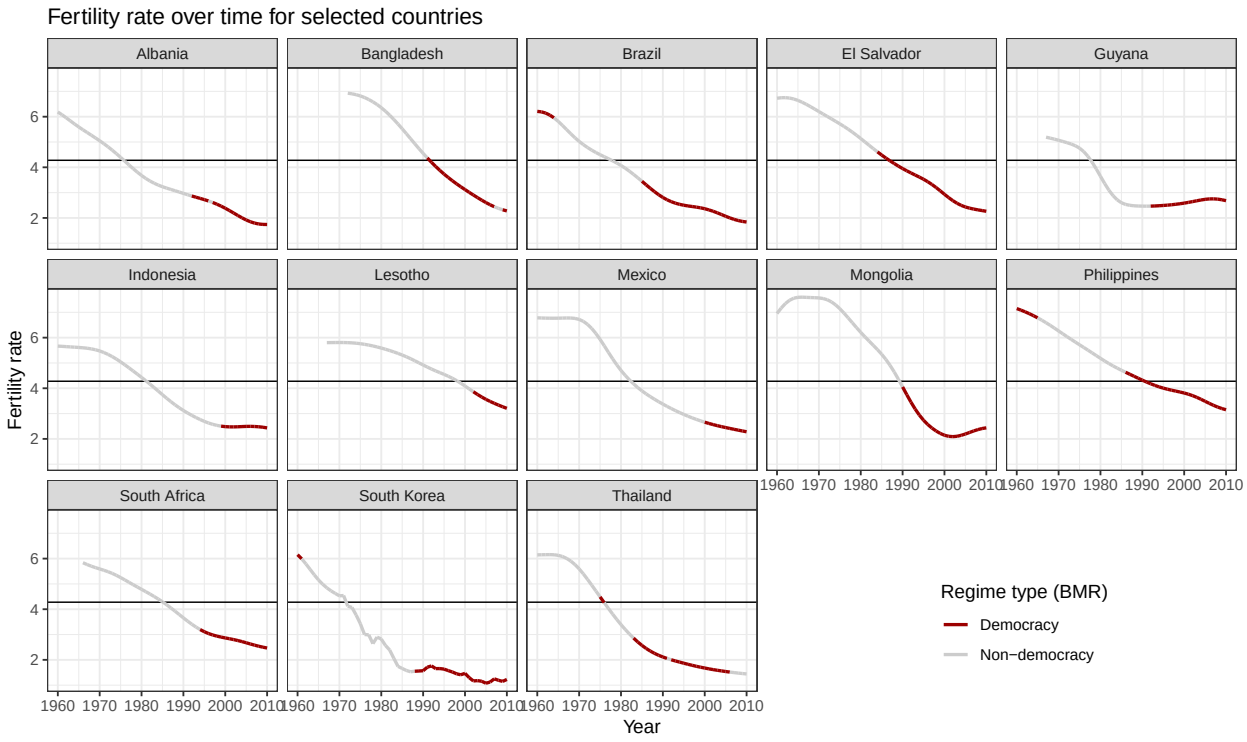


Figure 2: This figure shows fertility rate over time for selected countries during the observation period 1960-2010. We can see that these countries experienced democratization after their fertility rate plummeted below (or close to) the sample average (= 4.28, shown with the horizontal line in each panel).

fertility rate and conflict. One mechanism through which fertility rate could influence international conflict is through its effect on population size. For example, a high fertility rate may lead to the emergence of a youth bulge in the long term, which could, in turn, encourage states to be conflictual internally and externally (Collier 2000). Alternatively, a low fertility rate may imply a shrinking domestic market, requiring peaceful external trade relationship with foreign economies to secure an export market. Both of these arguments expect that fertility rate and conflict may be positively correlated due to a long-term demographic change.

Although empirical studies of international conflict typically do not consider demographic variables (such as population growth) as a relevant righthand-side variable, this potential oversight is hardly sufficient to support our claim that fertility rate does not have a direct influence on international conflict. One of the few systematic studies that

do look at demographic factors as a determinant of conflict behavior reports that population growth is (albeit modestly) positively associated with conflict involvement (Tir and Diehl 1998). As far as a declining fertility rate implies slower population growth and a decline of youth bulges,[27] this finding may cast doubt on the exogeneity of fertility rate as an instrument.

As pointed out by a later study, however, the effect of demographic pressure on conflict behavior is unlikely to be a direct one. More specifically, Cranmer and Siverson (2008) argue that autocratic leaders are not as responsive to population pressure as democratic leaders are and find evidence that population growth increases conflict initiation only among democracies. Building on their argument, we claim that the part of population pressure and market pressure associated with fertility rate will first influence regime type, which, in turn, will affect leaders' conflict behavior. We will address the issue of instrument validity more thoroughly by conducting a series of falsification tests that illustrate the plausibility of our assumptions. We will revisit this after we present the main empirical findings.

Other covariates

Although we claim that fertility rate is exogenous to militarized conflict, we acknowledge that fertility rate is likely to be endogenous to some economic, societal, and political forces. To reduce the possibility that such factors also influence conflict, opening up a backdoor path, we control for potential confounders of fertility rate in all three equations. First, we control for the level of economic development measured as per capita GDP. The expectation is that economic development reduces the necessity for families to have many children as workforce. Per capita income is also one of the strongest determinants of democracy identified in the literature as well (e.g., Przeworski, Alvarez, and Limongi 2000). As mentioned above, we expect that economic development also influences conflict, thus including it in the conflict equation as well.

[27] We note that this may not necessarily be the case. If a decline in fertility rate is accompanied by a decline

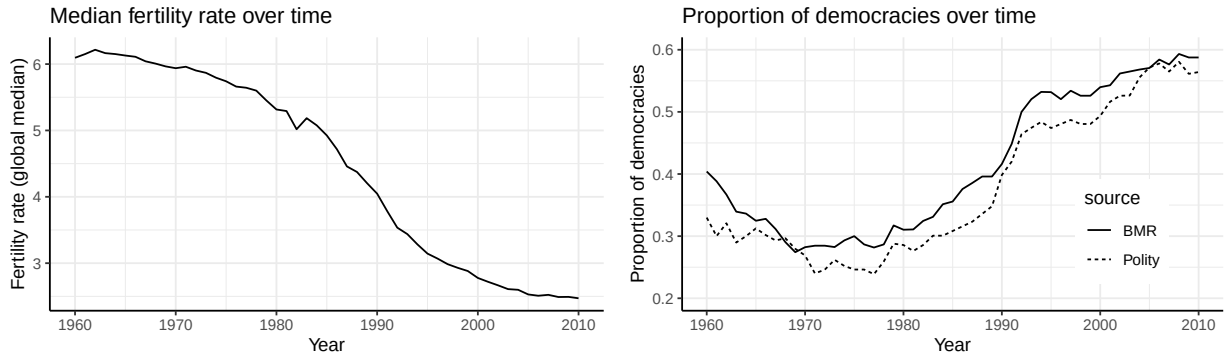


Figure 3: These figures show the secular trends of fertility rate (lefthand side) and the proportion of democracies in the world for the period between 1960 and 2010.

Another strong predictor of fertility rate is calendar year. As the lefthand side panel of Figure 3 shows, fertility rate has been generally declining over time for the observation period between 1960 and 2010. The righthand side panel of the same figure shows that, during the same time window, proportion of democratic countries in the world initially declines, starts increasing around year 1980, and then somewhat flats out from the 1990s on. To control for these secular trends, we include decades-level fixed effects in all three equations.

Main Findings

Table 2 shows the results of univariate and trivariate models of democracy and militarized conflict initiation. The first two numerical columns show the results for the BMR democracy measure, whereas the last two numerical columns show the results for the Polity democracy measure. For each of the two datasets with different democracy measures, we estimate a univariate clog-log model that ignores the endogeneity problem and a trivariate model that corrects for the endogeneity. By comparing models (1) and (2) or (3) and (4), we can see how our inferences change when we correct endogeneity, while holding everything else constant.

in infant mortality rate, their effects on population size may cancel each other out.

Table 2: Models of Democracy and Conflict Initiation, 1960–2010

	BMR		Polity	
	(1) Univariate	(2) Trivariate	(3) Univariate	(4) Trivariate
1: Conflict equation (clog-log)				
Challenger's Democracy	0.067	0.949***	0.259*	0.961***
	(0.139)	(0.297)	(0.143)	(0.301)
Target's Democracy	0.319**	0.388	0.391***	0.213
	(0.127)	(0.382)	(0.104)	(0.325)
Joint Democracy	−0.870***	−0.807***	−1.142***	−1.101***
	(0.213)	(0.196)	(0.209)	(0.192)
Military Advantage	0.560***	0.526***	0.632***	0.594***
	(0.131)	(0.131)	(0.142)	(0.148)
Challenger's per capita GDP (logged)	−0.044	−0.172***	−0.088*	−0.183***
	(0.049)	(0.059)	(0.051)	(0.060)
Target's per capita GDP (logged)	−0.069	−0.102*	−0.015	−0.008
	(0.049)	(0.059)	(0.049)	(0.052)
Peace Years (logged)	−0.993***	−0.973***	−0.967***	−0.960***
	(0.037)	(0.039)	(0.039)	(0.039)
Decade fixed-effects	✓	✓	✓	✓
2: Challenger's democracy equation (probit)				
Challenger's Fertility		−0.273***		−0.325***
		(0.022)		(0.024)
Challenger's per capita GDP (logged)		0.222***		0.226***
		(0.037)		(0.040)
Decade fixed-effects		✓		✓
3: Target's democracy equation (probit)				
Target's Fertility		−0.272***		−0.325***
		(0.022)		(0.024)
Target's per capita GDP (logged)		0.222***		0.227***
		(0.037)		(0.040)
Decade fixed-effects		✓		✓
Correlation coefficients				
ρ_{12}		−0.276***		−0.207***
		(0.068)		(0.077)
ρ_{13}		−0.079		0.012
		(0.085)		(0.072)
ρ_{23}		0.349***		0.339***
		(0.030)		(0.036)
Number of observations	42,072	42,072	35,940	35,940

$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$
Standard errors (in parenthes) are clustered at directed-dyad level. Estimates for constant in each of the three equations are omitted for brevity.

Let us first look at the results from the "first stage" of our instrumental variables estimation, where challenger's and target's democracy variables are regressed. Consistent with our expectation, the coefficients for the fertility rate variables are negative and highly statistically significant in both BMR and Polity models. Note, however, that variables in the "first-stage" models vary only across country and year but not across different directed-dyad-years for a given challenger (target) in a given year. This means the standard errors

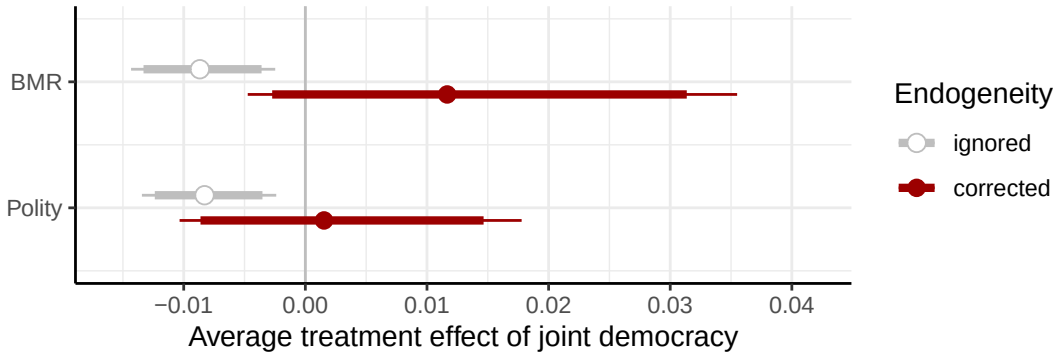


Figure 4: This figure shows the estimated average treatment effect of joint democracy on militarized conflict. Solid circles shown in red are the point estimates from the trivariate models that correct for the endogeneity problem, whereas hollow circles shown in gray are those from the univariate models that ignore the endogeneity. Horizontal line segments cover the 95% (thin) and 90% (thick) confidence intervals of the estimates.

for the first-stage equations may be underestimated. To guard against this possibility, we estimate univariate probit models of regime type using country-year as the unit of observation. [28] Fertility rate continues to be highly statistically significant ($p < 0.001$) in these country-year models. This provides strong support for the relevance of the instrument.

We now turn to the results from the outcome stage, where militarized conflict initiation is regressed on endogenous democracy measures and other covariates. Our univariate clog-log models successfully replicate the democracy–peace correlation, typically reported in previous studies. This illustrates the kind of inference we would make when we ignored the endogeneity of regime type. Specifically, joint democracy has a statistically significant negative coefficient, whereas individual democracy measures have either a positive or insignificant coefficient. As the sum of the three estimated coefficients for democracy measures [29] is negative, the results imply that jointly democratic dyads are less likely to experience militarized conflict even though individual democracies are not any less conflict-prone vis-a-vis non-democracies. We illustrate the substantive effects of joint democracy by calculating average treatment effects using equation (2). These effects are averaged over each of the observations in the estimation sample. [30] Hollow circles and

[28] Table A.4 in the Appendix shows the estimation results.
[29] This corresponds to $\hat{\beta}_1 + \hat{\beta}_2 + \hat{\beta}_3$ in equation (2).
[30] We obtain uncertainty estimates using the approach proposed by Krinsky and Robb (1986, 1990).

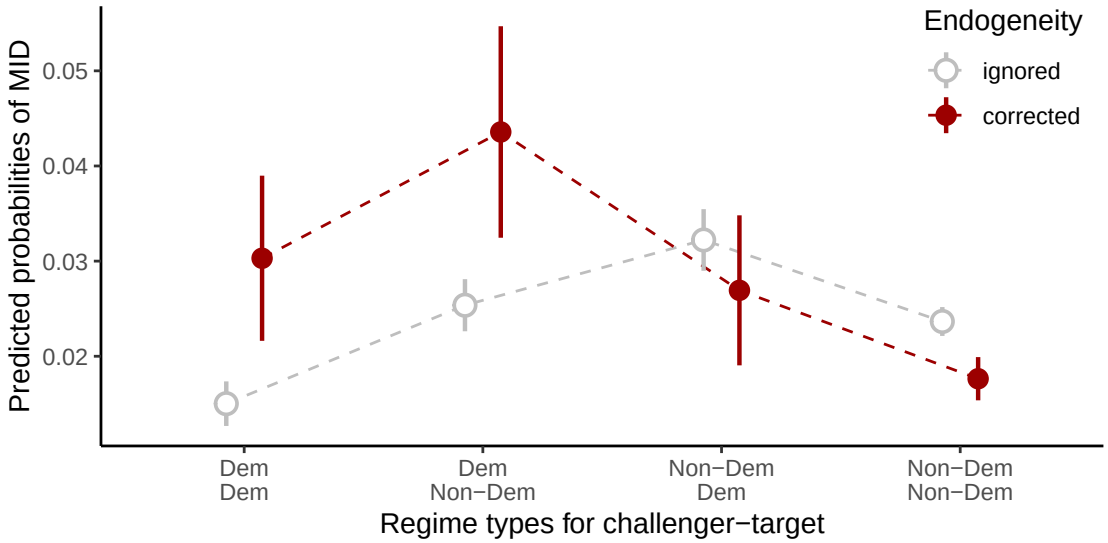


Figure 5: This figure shows the predicted probabilities of militarized conflict initiation for different combinations of regime types, based on the BMR data. Solid circles shown in red are the point estimates from the trivariate models, and hollow circles from the univariate models. Vertical line segments show the standard deviation of the predicted probabilities.

associated line segments in Figure 4 show the average effects of joint democracy estimated with the univariate clog-log models. The effects are negative and statistically significant at 95% confidence level for both BMR and Polity measures.

But, once we correct for the endogeneity of democracy using our trivariate model, this inference is no longer warranted. The estimated coefficients for joint democracy are still negative and statistically significant in both BMR and Polity models, but the coefficients for challenger's democracy are much greater than from the univariate models. The sum of the estimated coefficients for the three democracy measures is now positive for both BMR and Polity models, implying that the joint democracy may *increase* the probability of militarized conflict. Solid circles and associated line segments in Figure 4 show the estimated average effects of joint democracy from the trivariate models. We can see that the estimated effect of joint democracy is no longer negative for both BMR and Polity models. The point estimates are positive, although the estimates are not precise enough for us to reject the null hypothesis of no effect.

Whether we correct for endogeneity thus makes a significant difference in our esti-

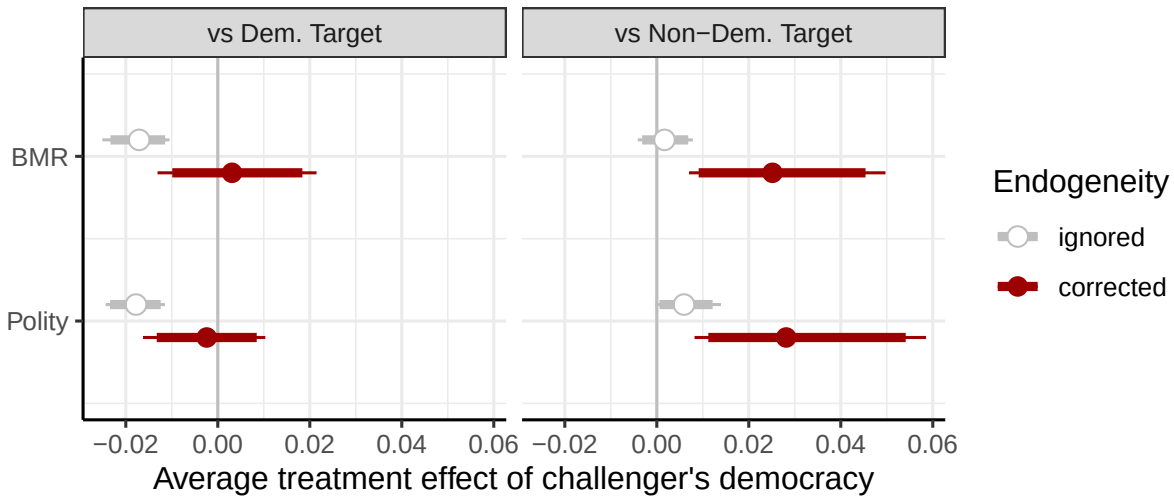


Figure 6: This figure shows the estimated average treatment effect of challenger's democracy on militarized conflict. These are calculated by holding constant the target's regime type at democracy (lefthand side panel) or non-democracy (righthand side panel).

mates of the effect of joint democracy on conflict. Let us now explore how this change comes about. Figure 5 shows the predicted probabilities of militarized conflict initiation for the four possible combinations of challenger's and target's regime types, estimated with the univariate (hollow circles) and the trivariate (solid circles) models. When we ignore endogeneity, the predicted probability of conflict is lowest for jointly democratic dyads. This is driven by two dynamics. First, when the target is democratic, a democratic challenger is less likely to initiate a conflict than a non-democratic challenger. [31] Second, when the target is a non-democracy, a democratic challenger is not more likely to initiate a conflict than a non-democratic challenger. [32]

These two dynamics no longer hold when we correct for endogeneity. First, when the target is democratic, challenger's regime type does not make a statistically significant difference. [33] Second, when the target is a non-democracy, a democratic challenger is more

[31] This is comparing the leftmost hollow circle (Dem/Dem) with the third hollow circle from left (Non-Dem/Dem) in Figure 5. Hollow circles in the lefthand side panel of Figure 6 illustrates this pacifying effect of challenger's democracy.

[32] This is comparing the rightmost hollow circle (Non-Dem/Non-Dem) with the second hollow circle from left (Dem/Non-Dem) in Figure 5. Hollow circles in the righthand side panel of Figure 6 illustrates this comparison.

[33] This is comparing the leftmost solid circle (Dem/Dem) with the third solid circle from left (Non-

likely to attack than would a non-democratic challenger. [34] The lefthand side panel of Figure 6 summarizes how the conflict-reducing effect of challenger's democracy (vs democratic target) disappears once we control for endogeneity. The righthand side panel of the same figure further shows that challenger's democracy encourages conflict (vs non-democratic target).

The key to understanding why these changes occur lies in the estimated correlation between the error terms for the conflict equation and the challenger's democracy equation, ρ_{12}. The estimated correlation coefficient is negative and statistically significant in both of the trivariate models, suggesting that unobservable factors that encourage a country's democracy make it less likely for that country to attack another country. A failure to control for such factors would generate a negative omitted variable bias, making it look that democracy has a pacifying effect.

Falsification Tests

The empirical findings we have presented so far critically depend on our identification assumptions, the most notable of which is the exclusion restriction — fertility rate must influence militarized conflict exclusively through regime type. This assumption would be violated if there is a causal link from fertility rate to conflict that does not go through regime type. Although this assumption cannot be verified empirically, it can still be falsified (Labrecque and Swanson 2018). We now assess the plausibility of the exclusion restriction by conducting a series of falsification tests.

[34] Dem/Dem) in Figure 5. Solid circles in the lefthand side panel of Figure 6 illustrates this comparison. This is comparing the rightmost solid circle (Non-Dem/Non-Dem) with the second solid circle from left (Dem/Non-Dem) in Figure 5. Solid circles in the righthand side panel of Figure 6 illustrates this conflict-inducing effect of challenger's democracy.

Test 1: Fertility rate and leader gender

We have argued above that a decline in fertility rate leads to female empowerment in a society, which encourages democratization and stabilizes an existing democracy. Female empowerment, however, may also have a direct impact on international conflict, and if so this violates the exclusion restriction. An emerging body of research has focused on the effect of gender on international relations (Reiter 2015). Most notably, seminal work by Caprioli (2000, 2003) suggest that gender equality measures, including fertility rate, influence international conflict. [35] We assess the extent to which this poses a threat to our identification strategy.

We begin by examining the theoretical mechanism linking fertility rate and conflict that Caprioli (2000, 2003) and other scholars have proposed. Their argument follows several steps. First, it is based on a premise that (i) women tend to have a more peaceful foreign policy preference than men. Second, it assumes that (ii) a declining fertility rate leads to female empowerment, which, in turn, contributes to gender equality. Third, (iii) as gender equality in a society improves, the state's foreign policy becomes more likely to reflect women's preferences.

We agree with the first premise, [36] but we doubt that the second and the third premises could operate independent of regime type. As we argue above, when gender equality improves thanks to female empowerment, the regime is very likely to become more democratic. In other words, the process anticipated by premise (ii) leads to democratization. Premise (iii) assumes that a country's foreign policy *automatically* reflects the preferences

[35] See also Regan and Paskeviciute (2003) and Sobek, Abouharb, and Ingram (2006). Regan and Paskeviciute (2003) find the opposite of what their theory expects about the relationship between fertility rate and conflict, and Sobek, Abouharb, and Ingram (2006) report mixed evidence for the relationship between respect for women's rights and conflict.

[36] Experimental studies have found somewhat mixed evidence for this. For example, Press, Sagan, and Valentino's (2013) experiment finds no significant difference between men and women in terms of their preference for the use of nuclear weapons. That said, there are some compelling biological and sociological theories that expect males to be more aggressive than females. For example, men tend to have higher levels of testosterone than women due to biological (McDermott et al. 2007) and non-biological (Klinesmith, Kasser, and McAndrew 2006) reasons, and this hormone is positively correlated with aggression.

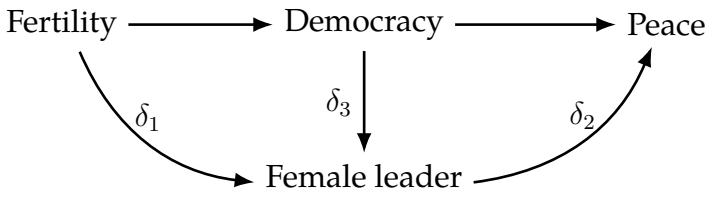


Figure 7: This diagram illustrates a scenario where the exclusion restriction may be violated. The exclusion restriction would be violated when both δ_1 and δ_2 are non-zero at the same time, because that opens up a path that connects fertility and peace that does not go through democracy. The exclusion restriction is still valid if the effect of fertility on leader gender goes through democracy rather than directly (i.e., $\delta_3 \neq 0$ and $\delta_1 = 0$).

of those constituents that are more powerful than other constituents in the society. This, however, would not happen in the absence of some mechanism that incentivizes the government to implement policies that reflect powerful constituents' preferences. Let us now consider what factors, other than democracy, could motivate the government to adopt foreign policies that reflect women's preferences more.

One possible way for empowered women to have an influence on foreign policy decision making *in the absence of democracy* would be to have a female national leader. When lower fertility rate empowers women, the probability that a woman becomes the national leader may increase, even in a non-democratic country. A female national leader may then implement foreign policies that reflect policy preferences of women. [37] Figure 7 presents a causal diagram that shows how this could violate the exclusion restriction. Suppose female empowerment due to a declining fertility rate has a direct impact on leader's gender (i.e., $\delta_1 \neq 0$). Suppose also that leader's gender has a direct influence on war and peace (i.e., $\delta_2 \neq 0$). [38] A non-zero δ_1 and a non-zero δ_2 together open up a back-door causal path from fertility to conflict that does not go through regime type, violating the exclusion restriction.

[37] Another causal pathway may be through nominal female representation in congress. Past studies have used percentage of women in the upper (Caprioli 2000) or the lower (Regan and Paskeviciute 2003) house of parliament as a measure of gender equality. As Bjarnegård and Melander (2013) point out, however, such measures are heavily influenced by gender quotas and may not accurately reflect the degree that political power is shared between men and women.

[38] That said, there is no consensus in the literature about whether leader gender has a direct impact on international conflict. For example, Horowitz, Stam, and Ellis (2015) find no evidence that leader gender influences conflict behavior.

Table 3: Clog-log Models of Female Leaders

	(5)	(6)	(7)	(8)
Fertility rate	−0.315*		−0.136	−0.379
	(0.174)		(0.202)	(0.425)
Democracy		2.759***	2.651***	1.416
		(0.671)	(0.714)	(1.685)
Fertility rate × Democracy				0.272
				(0.380)
N	7,345	7,345	7,345	7,345

*$p < 0.10$, **$p < 0.05$, ***$p < 0.01$
All models include the same set of control variables (logged per capita GDP, decade fixed effects, and controls for duration dependence) and the intercept. Estimated coefficients for these variables are suppressed for brevity. Standard errors (in parentheses) are clustered at country level.

If, on the other hand, any effect of fertility on leader gender *goes through* regime type (i.e., $\delta_3 \neq 0$) and not directly (i.e., $\delta_1 = 0$), the exclusion restriction is still valid. Our first falsification test thus examines whether fertility rate has a direct impact on the likelihood that a country has a female leader, independent of its effect that goes through regime type. We estimate a series of binary time-series cross-sectional clog-log models using country-year as the unit of observation. [39] The dependent variable is a dummy variable coded as 1 for a country that has at least one female national leader in a given year, and 0 otherwise. [40]

Table 3 reports the estimation results. Models (5) and (6) include fertility rate and regime type separately, whereas model (7) includes the two variables at the same time, and model (8) further includes an interaction term between the two. A lower fertility rate appears to increase the probability of a female leader when ignoring regime type (Model 5), but this relationship disappears when we control for regime type (Models 7 and 8). A plausible explanation for these findings is that any effect of fertility rate on leader gender goes exclusively through regime type, lending support for the exclusion restriction. [41]

[39] We use clog-log link function because having female leaders is a rare event (2.67% of all country-year observations have a female leader).

[40] We rely on information from version 4.1 of the Archigos data set (Goemans, Gleditsch, and Chiozza 2009).

[41] We also perform causal mediation analyses using Models (7) and (8) to estimate the direct effect of fertility rate on leader gender (δ_1) and the indirect effect (δ_3, the effect mediated by regime type). The

Test 2: Reduced-form analysis of conflict

Another way to assess the plausibility of the exclusion restriction is to explore reduced-form relationships between fertility rate and militarized conflict. [42] A reduced-form model omits the endogenous variable (i.e., regime type) from the outcome regression and looks at the direct relationship between the instruments (i.e., fertility rate) and the outcome (i.e., militarized conflict). That is, we simply replace the regime type variables with fertility variables in the univariate clog-log model predicting conflict. Given our IV estimates, we expect to find a significant negative relationship between challenger's fertility and conflict when we conduct a reduced-form analysis on our entire sample.[43]

But the exclusion restriction implies that a reduced-form analysis should reveal no relationship between fertility variables and conflict, if we are to analyze those cases where the causal path from fertility to democracy is blocked. In other words, the exclusion restriction would be violated if fertility rate is found to influence conflict in situations where the fertility-democracy path is shut out. Our second falsification test thus examines a reduced-form relationship for some subsets of our sample, for which it is plausible to assume that the path from fertility to regime type is blocked.

How can we create or find a situation where the path from fertility to conflict is plausibly blocked? We seek to simulate such an instance by considering who would be "always-takers" or "never-takers" if we were to conduct an ideal randomized control trial (RCT) on this topic. With regime type being our "treatment" variable (democracy and non-democracy are the possible values) and conflict being the outcome variable, fertility rate corresponds to the "treatment assignment" variable in an ideal RCT. For example, receiving a low fertility rate would correspond to it being assigned to the "treated" group (i.e.,

results (included in section E of the Online Appendix) suggest that the average direct effect is indistinguishable from zero, whereas average mediation effect is negative and significant. These findings also lend support to the exclusion restriction.

[42] The analysis below is motivated by Nunn and Wantchekon (2011) and Acharya, Blackwell, and Sen (2016).

[43] Challenger's fertility negatively influences democracy, which, in turn, should be positively correlated with conflict.

assigned to be democratic). Then, "always-takers" in this context correspond to those countries that would always be democratic regardless of their treatment assignment (i.e., regardless of fertility rate). Similarly, "never-takers" are the ones that would never be democratic regardless of their fertility rate. [44] By definition, the treatment assignment (fertility rate) should have no effect on the actual treatment (regime type) for always-takers and never-takers. Because the path from fertility to regime type is blocked for these subjects, reduced-form analyses of such cases should reveal no relationship between fertility rate and conflict — otherwise, there must be a direct path from fertility to conflict, violating the exclusion restriction.

Of course, it is impossible to know whether a particular country is an always-taker (never-taker) or a complier simply by observing the treatment assignment (i.e., fertility rate) and actual treatment status (i.e., regime type). [45] That being said, we argue we could narrow down the candidates by using information on treatment assignment and treatment status. Specifically, we classify democratic country-years with fertility rate higher than its median value as always-takers, and non-democratic country-years with fertility rate higher than its median value as never-takers. [46]

Figure 8 shows the estimated marginal effect of challenger's fertility rate on probability of militarized conflict for different values of target's fertility rate. As the lefthand side figure shows, when we analyze the reduced-form relationship for all observations, challenger's fertility rate has a negative effect on probability of conflict when the target's fertility rate is higher than about 5. However, this relationship disappears when we ana-

[44] An experimental subject could be either an always-taker, never-taker, complier, or defier. A complier takes treatment if and only if they are assigned to the treatment group. A definer takes treatment if and only if they are *not* assigned to the treatment group. Assuming there is no defier in our data, democratic countries are either a complier or always-taker. Non-democratic countries are either a complier or never-taker.

[45] For example, a country that has a low fertility rate and democratic regime could be either a complier or an always-taker. To find out which, we need to know what would happen in a counterfactual scenario where we hypothetically increased the fertility rate for the same country.

[46] For example, China has experienced a rapid decline in fertility rate due to the "one-child" policy since 1979, while it has always been non-democratic. According to our classification, China is coded as never-takers for years after 1975.

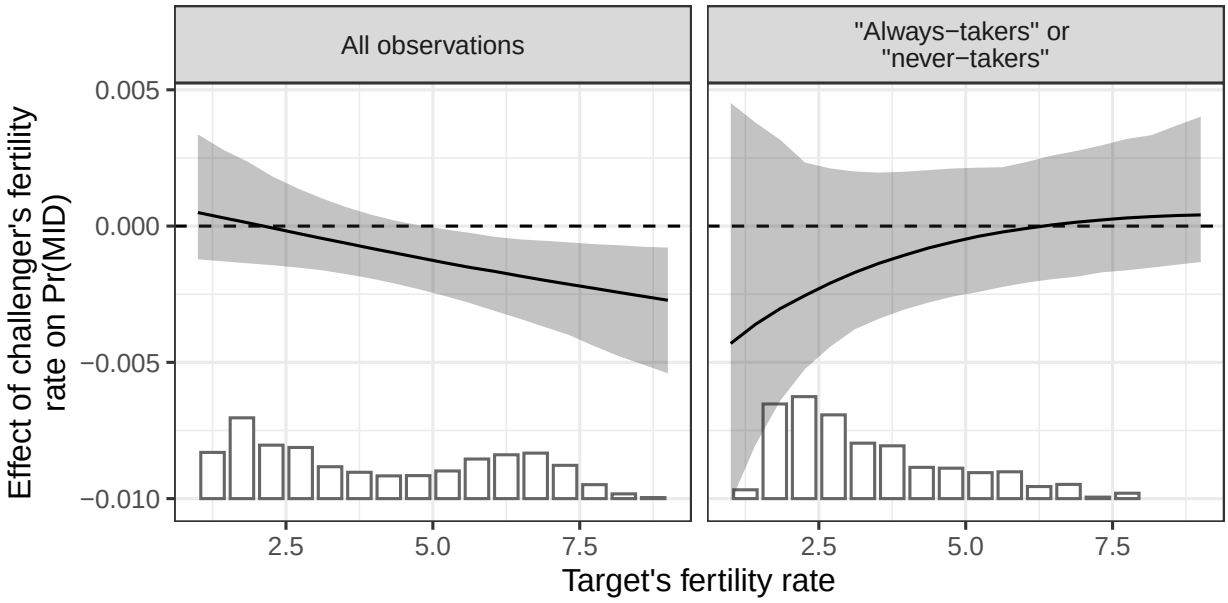


Figure 8: This figure shows the estimated marginal effect of challenger's fertility rate on the probability of militarized conflict. Marginal effect is calculated by changing the value of challenger's fertility rate by one standard deviation from its mean value, for different values of target's fertility rate. The lefthand side figure shows the results from the reduced-form model for all observations, whereas the righthand side figure is for a sample that includes only "always-takers" or "never-taker" as defined in the text.

lyze the sample that includes either always-takers or never-takers as defined above. These findings lend support to the exclusion restriction.

Test 3: Conflict and fertility

Another possible violation of the exclusion restriction could occur when fertility rate is influenced by militarized conflict. Specifically, past militarized conflict may influence present-day fertility rate, creating a correlation between fertility and conflict, independent of regime type. To guard against this possibility, our final falsification test examines the relationship between (change in) fertility rate as the dependent variable and past conflict involvement as independent variable.

Table 4 shows the results from linear regression models predicting (change in) fertility rate for country-year data. Models (9)–(11) use raw fertility rate as the dependent variable, whereas Models (12)–(14) use annual change in fertility rate as the dependent

Table 4: OLS models of fertility

	DV=fertility			DV=Δ fertility		
	(9)	(10)	(11)	(12)	(13)	(14)
# of conflicts in the previous year	−0.088			−0.003		
	(0.097)			(0.003)		
# of conflicts in the past 3 years		−0.044			−0.002	
		(0.045)			(0.002)	
# of conflicts in the past 5 years			−0.032			−0.002
			(0.031)			(0.001)
per capita GDP (logged)	−1.066***	−1.077***	−1.084***	0.003	0.004*	0.004*
	(0.085)	(0.083)	(0.082)	(0.002)	(0.002)	(0.002)
Constant	13.593***	13.631***	13.653***	−0.029	−0.038*	−0.046**
	(0.665)	(0.662)	(0.658)	(0.021)	(0.021)	(0.021)
Year, Year^2, Year^3	✓	✓	✓	✓	✓	✓
N	7,463	7,114	6,765	7,435	7,090	6,744

$^*p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$
Standard errors (in parentheses) are clustered at country level.

variable. We test three different measures of past conflict involvement; they count the number of militarized conflicts in the past 1 year, 3 years, or 5 years. In all the models, the coefficients for conflict count variables are indistinguishable from zero, lending support for the exclusion restriction. [47]

Conclusions

We propose an approach to estimate the causal effect of joint democracy on violent conflict between countries. Our empirical results suggest that democracy does not have a pacifying effect — if anything, democratic countries tend to attack other countries than non-democratic ones. These findings not only have important policy implications for foreign policy decision making but also contributes to the scholarly efforts to understand, explain, and predict state behaviors in world politics. Modern study of the democratic peace has

[47] These results are consistent with the findings from studies that look at the effect of *intrastate* conflict. For example, Urdal and Che (2013) find intrastate conflict both within and in the neighboring countries do not affect fertility rate.

started from an empirical observation that democracy and peace are correlated, which was then followed by theoretical efforts to make sense of the observed pattern. Given this trajectory, our finding that contradicts the initial empirical observation calls for a serious re-evaluation of many of the theories that have been proposed to explain the democratic peace.

If democracy is not the driving factor, we naturally wonder, what explains the observed peace between democracies? Unfortunately, answering this question is beyond the scope of this paper. That being said, future effort to explain this empirical pattern could build on our finding that it is the democratic challenger, not the democratic target, that is behind the increased risk of conflict between democracies. If some unobservable or unmeasured factors make it look that democratic pairs are more peaceful than other pairs, then the possibilities are: unobservable factors that encourage a country to be a democracy decreases the risk that it starts a militarized conflict, or unobservable factors that make a country militarily aggressive decreases the probability of it becoming a democracy. Identifying such a factor would be an obvious next step.

Another important avenue for future research is to examine the analytical procedures of our empirical work and assumptions underlying them to verify or, if necessary, modify our conclusions. As pointed out above, what our IV approach identifies is a *local* average treatment effect (LATE) — "local" in the sense that this is the fertility-induced causal effect of democracy, which may not be equal to the overall causal effect of democracy. What we show is that the part of the variation in democracy explained by the variation in fertility rate has a positive effect on conflict. It is still possible that the rest of the variation in democracy has a negative effect on conflict, but it is also possible that it has an even greater positive effect — we just cannot identify this effect without additional instruments and assumptions. We therefore agree with Imbens's (2010) aphorism: "better LATE than nothing." As we cannot do randomized controlled trial on democracy and peace, LATE is the best we could hope for in empirical research on the democratic peace. We hope that

other scholars follow our lead and conduct a study that identifies LATEs of joint democracy on conflict using alternative instruments.

References

Acemoglu, Daron, et al. 2019. "Democracy Does Cause Growth." *Journal of Political Economy* 127 (1): 47–100.

Acharya, Avidit, Matthew Blackwell, and Maya Sen. 2016. "The Political Legacy of American Slavery." *The Journal of Politics* 78 (3): 621–641.

Angrist, Joshua D., and Jorn-Steffen Pischke. 2009. *Mostly Harmless Econometrics*. Princeton University Press.

Beck, Nathaniel, and Jonathan N. Katz. 2001. "Throwing Out the Baby with the Bath Water: A Comment on Green, Kim, and Yoon." *International Organization* 55 (2): 487–495.

Bell, Mark S., and Kai Quek. 2018. "Authoritarian Public Opinion and the Democratic Peace." *International Organization* 72 (1): 227–242.

Betz, Timm, Scott J. Cook, and Florian M. Hollenbach. 2018. "On the Use and Abuse of Spatial Instruments." *Political Analysis* 26 (4): 474–479.

Bjarnegård, Elin, and Erik Melander. 2013. "Revisiting Representation: Communism, Women in Politics, and the Decline of Armed Conflict in East Asia." *International Interactions* 39 (4): 558–574.

Boix, Carles, Michael Miller, and Sebastian Rosato. 2013. "A Complete Data Set of Political Regimes, 18002007:" *Comparative Political Studies* 46 (12): 1523–1554.

Bueno de Mesquita, Bruce, et al. 1999. "An Institutional Explanation for the Democratic Peace." *American Political Science Review* 93 (4): 791–807.

Caprioli, Mary. 2003. "Gender Equality and State Aggression: The Impact of Domestic Gender Equality on State First Use of Force." *International Interactions* 29 (3): 195–214.

— . 2000. "Gendered Conflict." *Journal of Peace Research* 37 (1): 51–68.

Choi, Seung-Whan, and Patrick James. 2007. "Media Openness, Democracy and Militarized Interstate Disputes." *British Journal of Political Science* 37, no. 1 (): 23–46.

Choi, Sung-Whan. 2011. "Re-Evaluating Capitalist and Democratic Peace Models." *International Studies Quarterly* 55 (3): 759–769.

Colaresi, Michael P. 2014. *Democracy Declassified: The Secrecy Dilemma in National Security*. Oxford, New York: Oxford University Press.

Collier, Paul. 2000. "Doing Well out of War: An Economic Perspective." In *Greed and Grievance: Economic Agendas in Civil Wars*, ed. by David Malone, Mats R. Berdal, and International Peace Academy, 91–111. Boulder, Colorado: Lynne Rienner Publishers.

Cranmer, Skyler J., and Randolph M. Siverson. 2008. "Demography, Democracy and Disputes: The Search for the Elusive Relationship Between Population Growth and International Conflict." *The Journal of Politics* 70 (3): 794–806.

Dafoe, Allan. 2011. "Statistical Critiques of the Democratic Peace: Caveat Emptor." *American Journal of Political Science* 55 (2): 247–262.

Dafoe, Allan, John R. Oneal, and Bruce Russett. 2013. "The Democratic Peace: Weighing the Evidence and Cautious Inference." *International Studies Quarterly* 57 (1): 201–214.

Dyson, Tim. 2013. "On Demographic and Democratic Transitions." *Population and Development Review* 38 (s1): 83–102.

Farber, Henry S., and Joanne Gowa. 1997. "Common Interests or Common Polities?: Reinterpreting the Democratic Peace." *Journal of Politics* 59 (2): 393–417.

Fearon, James D. 1994. "Domestic Political Audiences and the Escalation of International Disputes." *American Political Science Review* 88 (3): 577–592.

Gartzke, Erik. 1998. "Kant We All Just Get Along? Opportunity, Willingness, and the Origins of the Democratic Peace." *American Journal of Political Science* 42 (1): 1–27.

— . 2000. "Preferences and the Democratic Peace." *International Studies Quarterly* 44 (2): 191–212.

— . 2007. "The Capitalist Peace." *American Journal of Political Science* 51 (1): 166–191.

Gartzke, Erik, and J. Joseph Hewitt. 2010. "International Crises and the Capitalist Peace." *International Interactions* 36 (2): 115–145.

Gibler, Douglas M. 2007. "Bordering on Peace: Democracy, Territorial Issues, and Conflict." *International Studies Quarterly* 51 (3): 509–532.

— . 2014. "Contiguous States, Stable Borders, and the Peace between Democracies." *International Studies Quarterly* 58, no. 1 (): 126–129.

— . 2012. *The Territorial Peace: Borders, State Development, and International Conflict*. Cambridge: Cambridge University Press.

Gibler, Douglas M., and Erin K. Little. 2017. "Heterogeneity in the Militarized Interstate Disputes (MIDs), 18162001: What Fatal MIDs Cannot Fix." *Political Science Research and Methods* 5, no. 1 (): 189–199.

Gibler, Douglas M., Steven V. Miller, and Erin K. Little. 2016. "An Analysis of the Militarized Interstate Dispute (MID) Dataset, 18162001." *International Studies Quarterly* 60, no. 4 (): 719–730.

— . 2020. "The Importance of Correct Measurement: A Response to Palmer, et Al." *International Studies Quarterly*.

Gibler, Douglas M., and Jaroslav Tir. 2014. "Territorial Peace and Democratic Clustering." *The Journal of Politics* 76 (1): 27–40.

Gleditsch, Kristian S. 2002a. "Expanded Trade and GDP Data." *Journal of Conflict Resolution* 46 (5): 712–724.

Gleditsch, Kristian Skrede. 2002b. *All International Politics Is Local: The Diffusion of Conflict, Integration, and Democratization*. University of Michigan Press.

Goemans, Henk E., Kristian Skrede Gleditsch, and Giacomo Chiozza. 2009. "Introducing Archigos: A Dataset of Political Leaders:" *Journal of Peace Research* 46 (2): 269–283.

Gowa, Joanne. 1999. *Ballots and Bullets: The Elusive Democratic Peace.* Princeton: Princeton University Press.

Green, Donald P, Soo Yeon Kim, and David H Yoon. 2001. "Dirty pool." *International Organization* 55 (2): 441–468.

Hegre, Håvard. 2000. "Development and the Liberal Peace: What Does it Take to Be a Trading State." *Journal of Peace Research* 37 (1): 5–30.

Horowitz, Michael C., Allan C. Stam, and Cali M. Ellis. 2015. *Why Leaders Fight*. Cambridge University Press.

Imbens, Guido W. 2010. "Better LATE Than Nothing: Some Comments on Deaton (2009) and Heckman and Urzua (2009)." *Journal of Economic Literature* 48 (2): 399–423.

Inglehart, Ronald, and Pippa Norris. 2003. *Rising Tide: Gender Equality and Cultural Change Around the World*. Cambridge, UK ; New York: Cambridge University Press.

Iversen, Torben, and Frances Rosenbluth. 2010. *Women, Work, and Politics: The Political Economy of Gender Inequality*. Yale University Press.

James, Patrick, Eric Solberg, and Murray Wolfson. 1999. "An Identified Systemic Model of the Democracy-peace Nexus." *Defence and Peace Economics* 10 (1): 1–37.

Johns, Robert, and Graeme A. M. Davies. 2012. "Democratic Peace or Clash of Civilizations? Target States and Support for War in Britain and the United States." *Journal of Politics* 74 (4): 1038–1052.

Klinesmith, Jennifer, Tim Kasser, and Francis T McAndrew. 2006. "Guns, Testosterone, and Aggression: an Experimental Test of a Mediational Hypothesis." *Psychological science* 17 (7): 568–571.

Krinsky, Itzhak, and A Leslie Robb. 1986. "On Approximating the Statistical Properties of Elasticities." *The Review of Economics and Statistics* 68 (4): 715–719.

— . 1990. "On Approximating the Statistical Properties of Elasticities: A Correction." *The Review of Economics and Statistics* 72 (1): 189–190.

Labrecque, Jeremy, and Sonja A. Swanson. 2018. "Understanding the Assumptions Underlying Instrumental Variable Analyses: A Brief Review of Falsification Strategies and Related Tools." *Current Epidemiology Reports* 5, no. 3 (): 214–220.

Maoz, Zeev, and Bruce Russett. 1993. "Normative and Structural Causes of Democratic Peace, 19461986." *American Political Science Review* 87 (3): 624–638.

Marshall, Monty G., Keith Jaggers, and Ted Robert Gurr. 2018. *Polity IV Project Dataset Users' Manual*.

McDermott, Rose, et al. 2007. "Testosterone and Aggression in a Simulated Crisis Game:" *The Annals of the American Academy of Political and Social Science* 614:15–33.

McDonald, Patrick J. 2015. "Great Powers, Hierarchy, and Endogenous Regimes: Rethinking the Domestic Causes of Peace." *International Organization* 69 (3): 557–588.

McGrath, Liam F. 2015. "Estimating Onsets of Binary Events in Panel Data." *Political Analysis* 23 (4): 534–549.

Mintz, Alex, and Nehemia Geva. 1993. "Why Don't Democracies Fight Each Other?: An Experimental Study." *Journal of Conflict Resolution* 37 (3): 484–503.

Mousseau, Michael. 2018. "Grasping the Scientific Evidence: The Contractualist Peace Supersedes the Democratic Peace." *Conflict Management and Peace Science* 35 (2): 175–192.

— . 2013. "The Democratic Peace Unraveled: It's the Economy." *International Studies Quarterly* 57 (1): 186–197.

Nunn, Nathan, and Leonard Wantchekon. 2011. "The Slave Trade and the Origins of Mistrust in Africa." *American Economic Review* 101 (7): 3221–3252.

Oneal, John R., and Bruce Russett. 1999. "Is the Liberal Peace Just an Artifact of Cold War Interests? Assessing Recent Critiques." *International Interactions* 25 (3): 213–241.

Paglayan, Agustina S. 2018. "Democracy and Educational Expansion: Evidence from 200 Years." *Unpublished manuscript* http://www.lse.ac.uk/lacc/publications/PDFs/Paglayan-Democracy-and-Educational-Expansion-130618.pdf.

Palmer, Glenn, et al. 2019. "Updating the Militarized Interstate Dispute Data: A Response to Gibler, Miller, and Little." *International Studies Quarterly*.

Park, Johann, and Michael Colaresi. 2014. "Safe Across the Border: The Continued Significance of the Democratic Peace When Controlling for Stable Borders." *International Studies Quarterly* 58, no. 1 (): 118–125.

Press, Daryl G., Scott D. Sagan, and Benjamin A. Valentino. 2013. "Atomic Aversion: Experimental Evidence on Taboos, Traditions, and the Non-Use of Nuclear Weapons." *American Political Science Review* 107 (1): 188–206.

Przeworski, Adam, Michael E. Alvarez, and Fernando Limongi. 2000. *Democracy and Development: Political Institutions and Well-Being in the World, 1950–1990*. Cambridge Studies in the Theory of Democracy. Cambridge University Press.

Rasler, K., and W. Thompson. 2005. *Puzzles of the Democratic Peace: Theory, Geopolitics and the Transformation of World Politics*. Evolutionary Processes in World Politics. Palgrave Macmillan US.

Reed, William, and Daina Chiba. 2010. "Decomposing the Relationship Between Contiguity and Militarized Conflict." *American Journal of Political Science* 54 (1): 61–73.

Regan, Patrick M., and Aida Paskeviciute. 2003. "Women's Access to Politics and Peaceful States." *Journal of Peace Research* 40 (3): 287–302.

Reiter, Dan. 2015. "The Positivist Study of Gender and International Relations." *Journal of Conflict Resolution* 59 (7): 1301–1326.

Reuveny, Rafael, and Quan Li. 2003. "The Joint DemocracyDyadic Conflict Nexus: A Simultaneous Equations Model." *International Studies Quarterly* 47 (3): 325–346.

Schultz, Kenneth. 2001. *Democracy and Coercive Diplomacy*. Cambridge University Press.

Schultz, Kenneth A. 1999. "Do Domestic Institutions Constrain or Inform?: Contrasting Two Institutional Perspectives on Democracy and War." *International Organization* 53 (2): 233–266.

— . 1998. "Domestic Opposition and Signaling in International Crises." *American Political Science Review* 94 (4): 829–844.

Sobek, David, M. Rodwan Abouharb, and Christopher G. Ingram. 2006. "The Human Rights Peace: How the Respect for Human Rights at Home Leads to Peace Abroad." *Journal of Politics* 68 (3): 519–529.

Sommer, Udi. 2018. "Women, Demography, and Politics: How Lower Fertility Rates Lead to Democracy." *Demography* 55 (2): 559–586.

Stinnett, Douglas M., et al. 2002. "The Correlates of War (Cow) Project Direct Contiguity Data, Version 3.0." *Conflict Management and Peace Science* 19 (2): 59–67.

Thompson, William R. 1996. "Democracy and Peace: Putting the Cart before the Horse?" *International Organization* 50 (1): 141–174.

Tir, Jaroslav, and Paul F. Diehl. 1998. "Demographic Pressure and Interstate Conflict: Linking Population Growth and Density to Militarized Disputes and Wars, 1930-89." *Journal of Peace Research* 35 (3): 319–339.

Tomz, Michael R., and Jessica L. P. Weeks. 2013. "Public Opinion and the Democratic Peace." *American Political Science Review* 107 (4): 849–865.

Urdal, Henrik, and Chi Primus Che. 2013. "War and Gender Inequalities in Health: The Impact of Armed Conflict on Fertility and Maternal Mortality." *International Interactions* 39 (4): 489–510.

Wilson, Ben, and Tim Dyson. 2017. "Democracy and the Demographic Transition." *Democratization* 24 (4): 594–612.

Make Two Democracies and Call Me in the Morning: Endogenous Regime Type and the Democratic Peace

Web Appendix

Not for Print Publication

A Monte Carlo Simulations

We assume the following data generation process. First, we generate a set of four covariates, x_{1i}, x_{2i}, x_{3i}, and x_{4i}, for $i = 1, 2, 3, \ldots, 1000$, each based on independent uniform distribution $\mathcal{U}(0, 1)$. We fix the values of these covariates constant throughout the entire simulation. For each round of simulation, we draw a set of three error terms, ϵ_{yi}, ϵ_{ci}, and ϵ_{ti}, from a trivariate Normal distribution. We fix the value of ρ_{23} (error correlated between challenger's and target's democracy) at 0.3 throughout the simulation, but change the value of ρ_{12} (error correlation between conflict and challenger's democracy) at $\{0, -0.2\}$ and the value of ρ_{13} (error correlation between conflict and target's democracy) at $\{0, -0.2, -0.4, -0.6\}$. Permutations of two different values of ρ_{12} and four different values of ρ_{13} produce eight combinations of the error correlations. For example, a combination $(\rho_{12} = 0, \rho_{13} = -0.4, \rho_{23} = 0.3)$ means that unobservable factors that encourage democracy are uncorrelated with unobservable determinants of challenger's conflict initiation but negatively correlated with unobservable determinants of target's propensity to attract conflict. Note that a combination $(\rho_{12} = 0, \rho_{13} = 0, \rho_{23} = 0.3)$ corresponds to a case where unobservable determinants of democracy are uncorrelated with conflict, suggesting that there is no endogeneity in a univariate estimation of equation (1).

For each of the eight combinations of ρ's, we perform 200 iterations of simulations. For each iteration, we generate endogenous variables D_i^c, D_i^t, and y_i as follows:

$$D_i^c = 1[2x_1 - x_3 + \epsilon_{ci} > 0] \tag{6}$$

$$D_i^t = 1[2x_2 - x_3 + \epsilon_{ti} > 0] \tag{7}$$

$$y_i = 1[-1 + \beta_1 D_i^c + \beta_2 D_i^t + \beta_3 D_i^c D_i^t + 0.5x_3 - 0.5x_4 + \epsilon_i^y > 0]. \tag{8}$$

We set $\beta_1 = 0.5$, $\beta_2 = 0.5$, and $\beta_3 = -0.5$, so that $\beta_1 + \beta_2 + \beta_3 = 0.5$. This means that

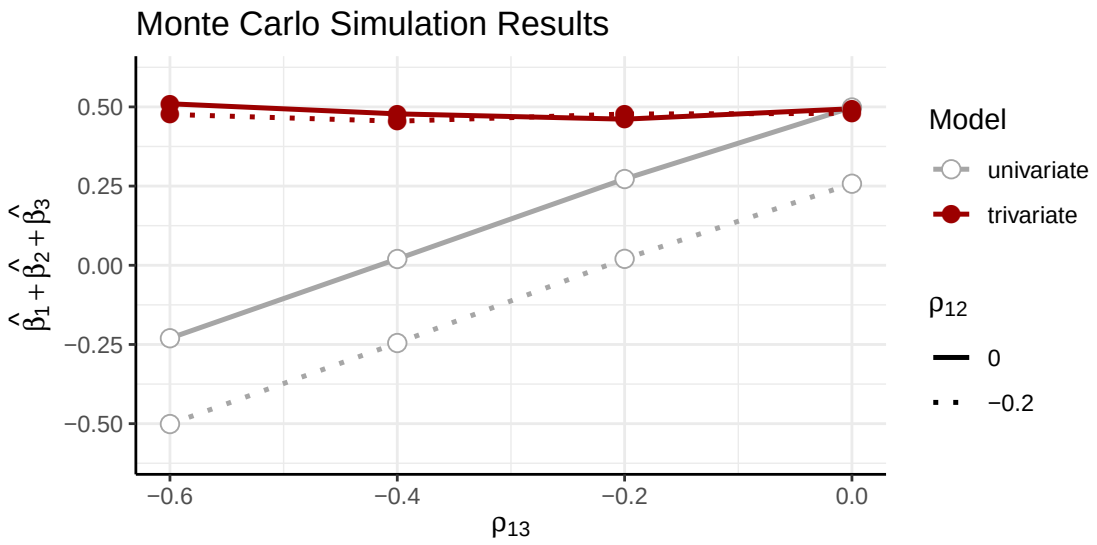


Figure A.1: This figure shows the results of Monte Carlo simulations. Solid circles show the mean values of the estimated $\beta_1 + \beta_2 + \beta_3$ from the trivariate probit models for different values of ρ_{12} and ρ_{13}, whereas hollow circles (in gray) are from the univariate probit models. Solid lines indicate ρ_{12} is set to 0, whereas dotted lines indicate ρ_{12} is set to -0.2. The true value of $\beta_1 + \beta_2 + \beta_3$ is held constant at 0.5 and ρ_{23} is held at 0.3 throughout the simulations.

joint democracy does *not* have a pacifying effect in truth. For each iteration, we estimate a univariate and a trivariate probit models and store the estimated values of $\hat{\beta}_1 + \hat{\beta}_2 + \hat{\beta}_3$ from each model.

Figure A.1 summarizes the results of our Monte Carlo simulations. As expected, when $\rho_{12} = 0$ and $\rho_{13} = 0$, both univariate and trivariate models can recover the true joint effect of democracy ($\beta_1 + \beta_2 + \beta_3 = 0.5$). When either ρ_{12} or ρ_{13} deviates from 0, however, univariate probit models produce bias while trivariate probit models consistently recover the correct value of $\beta_1 + \beta_2 + \beta_3$. We can see that, as ρ_{13} gets smaller, univariate estimation yields greater negative bias, making it more likely that we mistakenly conclude joint democracy has a pacifying effect.

B Comparison of Democracy Measures

Table A.1 lists all country-year observations that are coded as democracy by BMR but non-democracy by Polity, whereas table A.2 lists all country-year observations that are coded

as non-democracy by BMR but democracy by Polity.

Table A.1: List of country-years coded as democracy by BMR but non-democracy by Polity

Country	**Years**	**BMR**	**Polity**	**mean(Polity)**
Albania	1992–2001	Democracy	Non-democracy	5
Argentina	1960–61, 63–65	Democracy	Non-democracy	−1
Bangladesh	2009–10	Democracy	Non-democracy	5
Bolivia	1979	Democracy	Non-democracy	−4
Bosnia & Herzegovina	1993–94	Democracy	Non-democracy	0
Brazil	1961–63	Democracy	Non-democracy	4.33
Central African Republic	1993–2002	Democracy	Non-democracy	5
Chile	1960–63	Democracy	Non-democracy	5
Congo (Brazzaville)	1961–62	Democracy	Non-democracy	4
Dominican Republic	1966–77, 1994-95	Democracy	Non-democracy	−1.86
Ecuador	1960–62, 2007–10	Democracy	Non-democracy	2.86
France	1960–68	Democracy	Non-democracy	5
Ghana	1970–71, 97–2000	Democracy	Non-democracy	2.33
Greece	1960–66, 1974	Democracy	Non-democracy	3.62
Guatemala	1960–1962, 66–81, 86–95	Democracy	Non-democracy	−0.03
Guinea-Bissau	1994–97	Democracy	Non-democracy	5
Honduras	1960–62, 71, 85–88	Democracy	Non-democracy	2.12
Lebanon	1971–75	Democracy	Non-democracy	4
Malawi	2001–03	Democracy	Non-democracy	4.33
Moldova	1991–92	Democracy	Non-democracy	5
Mongolia	1990–91	Democracy	Non-democracy	2
Mozambique	1994–2003	Democracy	Non-democracy	5
Nepal	1991–98	Democracy	Non-democracy	5
Nicaragua	1984–89	Democracy	Non-democracy	−1
Niger	1999–2003	Democracy	Non-democracy	5
Pakistan	1972, 2008–09	Democracy	Non-democracy	4.67
Panama	1960–67	Democracy	Non-democracy	4
Papua New Guinea	1976–2010	Democracy	Non-democracy	4
Peru	1960–61, 63–67	Democracy	Non-democracy	5
Philippines	1960–64, 86	Democracy	Non-democracy	4.33
Poland	1989–90	Democracy	Non-democracy	5
Romania	1991–95	Democracy	Non-democracy	5
Russia	1992–98	Democracy	Non-democracy	3.29
Sierra Leone	2002–06	Democracy	Non-democracy	5
Spain	1977	Democracy	Non-democracy	5
Sri Lanka	1991–2000, 2003–05, 09	Democracy	Non-democracy	5
Suriname	1975–79, 88–89, 91–2010	Democracy	Non-democracy	4.56
Thailand	1975, 83–90	Democracy	Non-democracy	2.44
Turkey	1971–72	Democracy	Non-democracy	−2
Uganda	1980–84	Democracy	Non-democracy	3
Ukraine	1993	Democracy	Non-democracy	5
Uruguay	1971–72	Democracy	Non-democracy	0

Table A.2: List of country-years coded as democracy by Polity but non-democracy by BMR

Country	**Years**	**BMR**	**Polity**	**mean(Polity)**
Armenia	1992–94	Non-democracy	Democracy	7
Bangladesh	1972–73	Non-democracy	Democracy	8
Belarus	1994	Non-democracy	Democracy	7
Chile	1989	Non-democracy	Democracy	8
Comoros	2004–05	Non-democracy	Democracy	6
Cyprus	1960–62, 68–76	Non-democracy	Democracy	8
Dominican Republic	1962	Non-democracy	Democracy	8
Ecuador	2000–02	Non-democracy	Democracy	6
Gambia	1965–71	Non-democracy	Democracy	8
Guinea-Bissau	2005–10	Non-democracy	Democracy	6
Haiti	1990, 94–98	Non-democracy	Democracy	7
Honduras	2009	Non-democracy	Democracy	7
Latvia	1991–92	Non-democracy	Democracy	8
Lebanon	2005–10	Non-democracy	Democracy	6
Lesotho	1967–69, 93–97, 2001	Non-democracy	Democracy	8.11
Lithuania	1991	Non-democracy	Democracy	10
Madagascar	1992	Non-democracy	Democracy	9
Malaysia	1960–68, 2008–10	Non-democracy	Democracy	9
Mexico	1997–99	Non-democracy	Democracy	6
Namibia	1991–2010	Non-democracy	Democracy	6
Nepal	2006–07	Non-democracy	Democracy	6
Niger	1992	Non-democracy	Democracy	8
Nigeria	1983	Non-democracy	Democracy	7
Panama	1989–90	Non-democracy	Democracy	8
Paraguay	1992–2002	Non-democracy	Democracy	6.91
Peru	1990–91	Non-democracy	Democracy	8
Russia	2000–06	Non-democracy	Democracy	6
Solomon Islands	2004–05	Non-democracy	Democracy	8
South Africa	1992–93	Non-democracy	Democracy	7
Sri Lanka	1977–81	Non-democracy	Democracy	6.4
Turkey	1960	Non-democracy	Democracy	7
Uganda	1962–65	Non-democracy	Democracy	7
Venezuela	2005	Non-democracy	Democracy	6
Zambia	1991–95	Non-democracy	Democracy	6

Table A.3:

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Dispute initiation	42,072	0.026	0.159	0	0	0	1
Democracy	42,072	0.436	0.496	0	0	1	1
Joint democracy	42,072	0.290	0.454	0	0	1	1
Capability balance	42,072	0.500	0.337	0.00005	0.173	0.827	1.000
Per capita GDP (logged)	42,072	8.404	1.233	4.889	7.407	9.371	13.357
Fertility rate	42,072	4.075	2.097	1.076	2.067	6.073	8.864
Peace years	42,072	20.190	15.977	1	7	30	65

C Descriptive Statistics

D First-stage probit models

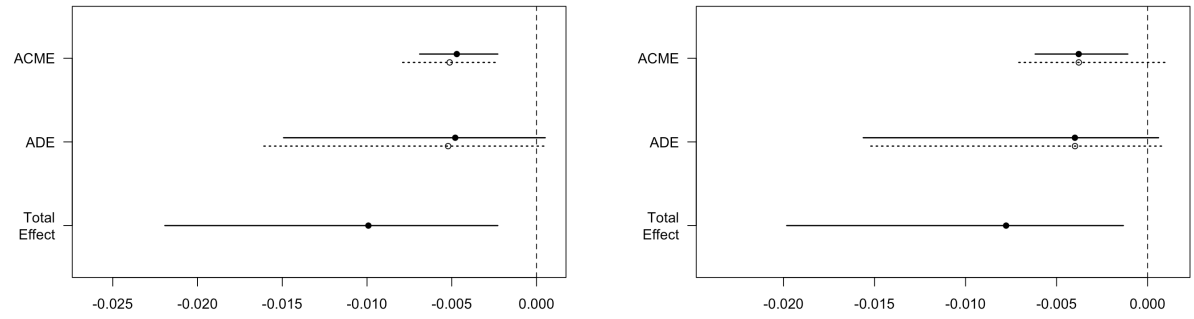
Table A.4: First-stage Models of Democracy, 1960–2010

	BMR		Polity	
	(1) with fertility	(2) w/out	(3) with fertility	(4) w/out
Challenger's Fertility	−0.297***		−0.325***	
	(0.055)		(0.057)	
Challenger's per capita GDP (logged)	0.175*	0.466***	0.209**	0.508***
	(0.093)	(0.085)	(0.098)	(0.040)
Decade fixed-effects	✓	✓	✓	✓
Number of observations	7,631	7,631	6,937	6,937
AIC	8,403.63	8,709.08	6,839.23	7,486.278
Area under ROC	0.81	0.77	0.83	0.79

$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$
Standard errors (in parenthes) are clustered at directed-dyad level. Estimates for constant in each of the three equations are omitted for brevity.

E Causal Mediation Analyses

We perform causal mediation analyses to estimate the *direct* effect of fertility on leader gender and the *indirect* effect of fertility on leader gender that goes through regime type. The first analysis is based on Model (7) in Table 3, which has leader gender (female = 1, male = 0) as the dependent variable and fertility rate and regime type along with other



Model (7) Model (8)

Figure A.2: These figures show the results of causal mediation analyses. The lefthand side panel is based on Model (7) and the righthand side panel is based on Model (8). ACMEs correspond to the indirect effect of fertility on leader gender that goes through regime type, whereas ADEs correspond to the direct effect of fertility on leader gender. Solid circles show the point estimates for democracies, and hollow circles for non-democracies. Line segments show 95% confidence intervals.

controls as the independent variables. The second analysis is based on Model (8), which adds an interaction term between fertility rate and regime type to Model (7). Our goal is to separate average causal mediation effect (ACME) of fertility rate (δ_3) from average direct effect (ADE) of fertility rate (δ_1). Results shown in Figure A.2 show that direct effects (δ_1, ADE) are indistinguishable from zero in both Models (7) and (8), suggesting that fertility does not have a direct effect on leader gender— its effect goes through (i.e., mediated by) regime type.