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

July 23, 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: 11,511 words (without appendix)

*We thank Nils-Christian Bormann, Kevin A. Clarke, Charles Crabtree, Shohei Doi, Songying Fang, Kentaro Fukumoto, Kristian Skrede Gleditsch, Håvard Hegre, Tobias Heinrich, Jesse C. Johnson, Azusa Katagiri, Kyosuke Kikuta, Atsushi Tago, Tore Wig, 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 in international relations, stimulating a lot of scholarly and policy debates for the past decades. Although a general consensus has emerged among scholars as to the existence of 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. Once we correct for this endogeneity via an instrumental variables approach, we find that joint democracy does not have a pacifying effect. On the contrary, our findings show that democratic countries are more likely to attack other democracies than non-democratic countries are. Our results call for a serious reevaluation of the theories proposed to explain the observed peace between democracies.

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 research design and data. After presenting the main findings, we report the results of a series of falsification tests to assess the plausibility of the assumptions behind our empirical strategy. 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 He-

(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.

witt 2010),[1] contractualist economy (Mousseau 2013, 2018),[2] stable territorial borders (Gibler 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

[1] See also Choi (2011) and Dafoe (2011).
[2] See also Dafoe, Oneal, and Russett (2013).
[3] See also Park and Colaresi (2014) and Gibler (2014).

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 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 an explanation for the democratic peace, it alone is insufficient. Studies of public opinion are an indirect test and are not tests of the causal relationship between democracy and war. The relationship between preference of the public and observed behavior of a country could be more complicated due to strategic interaction. For example, if public prefer to avoid fighting with democ-

[4] Note that a later study provides an important caveat to this finding. 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. This suggests that the difference in public opinion cannot be the cause of the difference in conflict propensity between democracy-democracy pairs and other pairs.

racies, this may embolden democracies in crisis bargaining, which may in turn lead to an increase in the 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 between democracy and peace, using the commonly used observational data. Such an approach allows us to study the overall relationship between regime type and war (rather than intermediate outcomes such as levels of public support for war) 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. [5] 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). [6]

[5] 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.

[6] 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 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.

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 war. 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. [7]

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 lower fertility rate is a strong driver of democratization. We will also present theoretical arguments and a series of falsification tests that support the claim

[7] 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 the opposite to justify his instruments.

that fertility rate does not directly influence international conflict.

Research Design

We analyze observational data on democracy and conflict for the period between 1960 and 2010. [8] Our sample includes all contiguous directed dyad years identified by the Correlates of War (COW) data. A directed dyad is a country pair 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 [9]) to reduce the heterogeneity in the sample and make each observation unit more comparable. [10] 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 should thus enhance the internal validity of our study, although it may come at a cost of somewhat diminished external validity. [11]

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. [12] To reduce heterogeneity

[8] 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.

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

[10] 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.

[11] To assess the generalizability of our findings beyond contiguous dyads, we expand the scope by including all country-pairs within the same geographic region as well. This makes the sample size more than five times bigger, from 34,312 to 199,168, and the results are unchanged. The results are included in the Online Appendix.

[12] 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).

in the MID data, we omit "protest-dependent" MIDs that are based on a substantially different coding logic (Gibler and Little 2017). [13] Our outcome variable, conflict initiation, 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 code regime type using the combined democracy-autocracy scale (polity2 variable) from the Polity5 Project (Marshall, Jaggers, and Gurr 2018). [14] Following the convention, we code a country as a democracy if it receives the polity2 score of 6 or higher. [15] To assess the effects of regime type on conflict, we regress our dependent variable, MID initiation, on challenger's regime type, target's regime type, and their interaction term. Including the interaction term, Joint Democracy, allows us to capture the dyadic nature of the democratic peace — democracies are expected to be peaceful only toward other democracies.

In analyzing the relationship between regime types and conflict, we acknowledge that some unobservable variables can influence these variables jointly, causing an endogeneity bias. To facilitate causal inference in face of this endogeneity, we employ a trivariate model that jointly estimates three equations for three endogenous binary variables (conflict, challenger's regime type, and target's regime type) as well as the correlations among the three error terms. The model is an application of a copula-based joint modeling framework for causal inference proposed by Braumoeller et al. (2018). [16] Our approach to model democracies and conflict jointly can be understood as an implementation of instrumental variables (IV) regression. [17]

In estimating the conflict equation of the model, we control for factors that could influ-

[13] An alternative approach to reduce heterogeneity of MIDs is to focus on those MIDs that involve fatalities (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.

[14] The data are available at `http://www.systemicpeace.org/polityproject.html` (Last accessed on July 16, 2020).

[15] We check the robustness of the findings by using alternative thresholds (5 and 7) to code democracy, and the results are unchanged. See the Appendix for the results using different thresholds.

[16] We provide a more detailed description of the model along with Monte Carlo simulation results in the Online Appendix.

[17] 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 generally equal to ATE, this is the best we can hope for when we cannot perform treatment randomization. We will revisit this issue later.

ence the probability that a challenger initiates conflict against a target. [18] 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 be more likely to have other means to compel other states or to deter aggression, economic development will influence the likelihood that a country experiences a militarized conflict either as a challenger or a target. [19] We thus control for both challenger's and target's economic development measured by per capita GDP. [20] 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. [21]

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, the two equations explaining democracy for challenger and target must include a significant determinant of democracy that does not belong in the equation 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 think of plausible theoretical mechanisms linking them with conflict.

We argue that fertility rate (the average number of children that a woman would have

[18] To reduce the danger of post-treatment variable bias, we opted not to include some of the well-known correlates of MIDs that may be influenced by regime type. For example, foreign policy similarity and dyadic trade may both be influenced by regime type and are thus excluded.

[19] As regime type may influence economic development, which would result in post-treatment bias, we also estimate a model that excludes economic development from all equations as a robustness check.

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

[21] We chose the natural log of time over other smooth functions (i.e., cubic polynomial, square root, and penalized thin plate regression splines) based on model fit statistics.

over her lifetime) in a country serves as a valid instrumental variable in estimating the effect of democracy on conflict. We start by presenting theoretical argument and empirical evidence that this is a relevant instrument. After presenting our argument for instrument exogeneity, we 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. [22] 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. [23] 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 [24] — 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 as-

[22] We acknowledge that it is difficult to completely rule out the possibility of reverse causality in the fertility-democracy relationship. Perhaps the relationship may be bi-directional — a lower fertility encourages democracy, and democracy lowers fertility rate. It is important to note, however, that this does not pose any problem for the purpose of obtaining a consistent estimate of the causal effect of democracy. For an instrument to be relevant, we need a correlation, not causation, between fertility and democracy.

[23] 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).

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

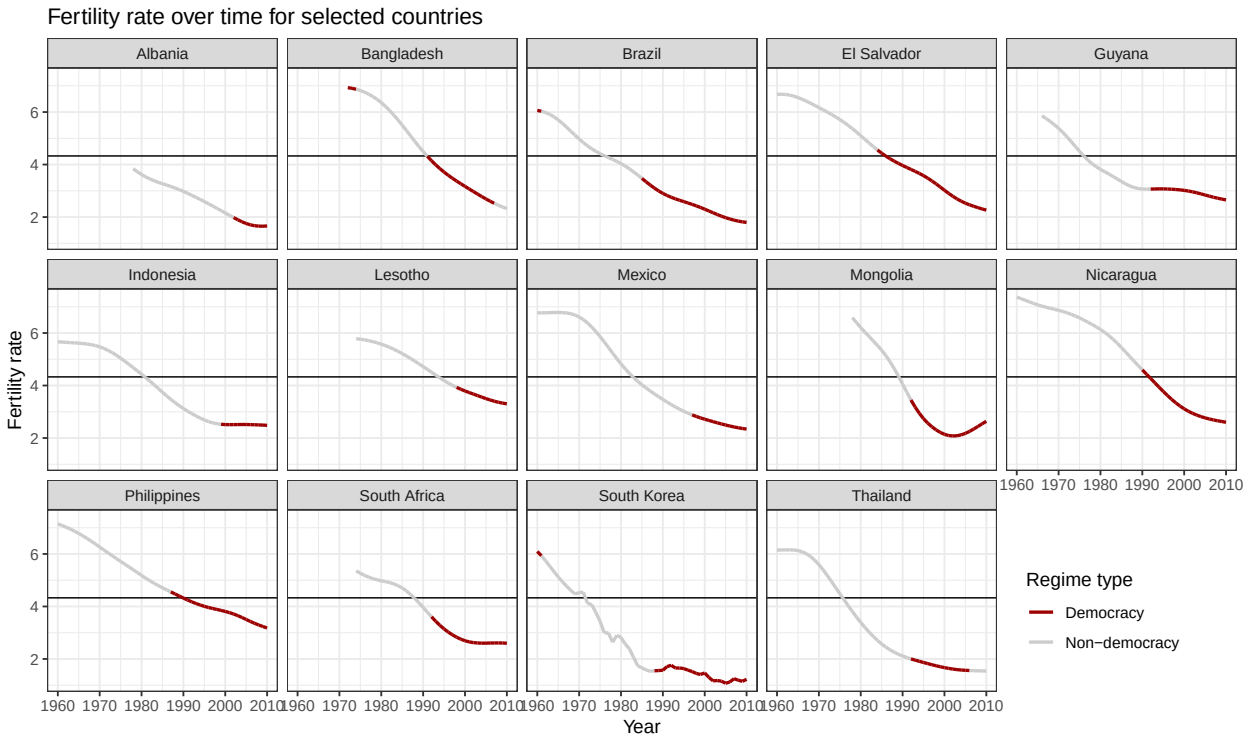

Figure 2: This figure shows fertility rate over time for selected countries during the observation period 1960-2010. We can see that these countries experienced democratization after their fertility rate plummeted below (or close to) the sample average (= 4.31, shown with the horizontal line in each panel).

sess its plausibility by contemplating on possible direct theoretical connections between 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,[25] 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 exogeneity 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, which would open up a backdoor path, we control for strong determinants of fertility rate in all three equations. [26] First, we control for infant mortality rates (log transformed) measured as the number

[25] We note that this may not necessarily be the case. If a decline in fertility rate is accompanied by a decline in infant mortality rate, their effects on population size may cancel each other out.

[26] Note that these explanatory variables we discuss below are intended to make fertility rate conditionally exogenous. We therefore include these variables not only in the democracy equations but also in the conflict equations, even though we may not have a clear expectation as to how these variables influence conflict.

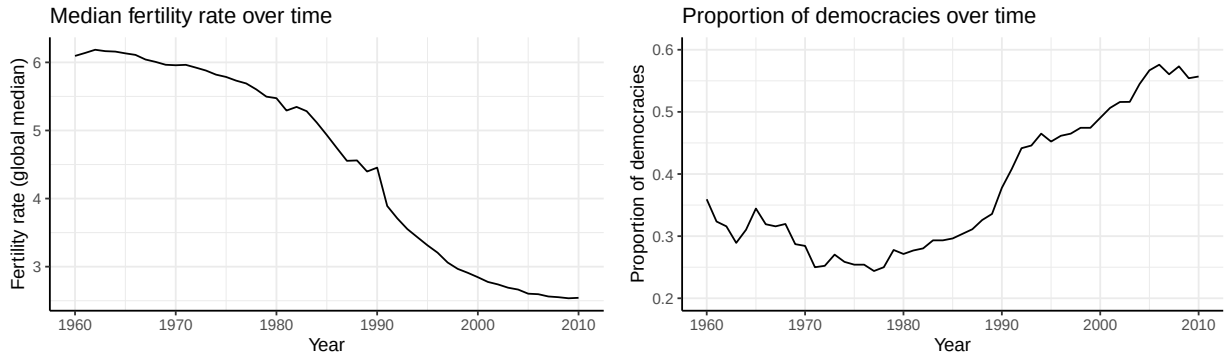

Figure 3: These figures show the secular trends of fertility rate (lefthand side) and the proportion of democracies in the world for the period between 1960 and 2010.

of infants dying before reaching one year of age, per 1,000 live births in a given year. [27] The expectation is that a lower infant mortality rate reduces the necessity for families to have many children as an insurance policy to secure workforce. Second, we control for the percentage of population living in urban areas, as a proxy for economic modernization and industrialization, which would improve access to modern contraceptives. [28] Urbanization would also enhance the costs of raising children due to higher housing prices of cities (vis-a-vis rural areas) and higher opportunity costs. Third, we include per capita GDP to account for any residual effect of economic development levels that is not captured by infant mortality and urbanization. [29]

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. During the same time window, proportion of democratic countries in the world has been increasing, albeit at a varying pace, as shown in the righthand side panel of the same figure. To guard against the possibility that unmeasured secular trend and regional idiosyncrasies influence both fertility, democracy, and conflict, we control for decade and region fixed-effects in all three equations.

[27] Data on mortality rate are obtained from the World Bank, available at: https://data.worldbank.org/indicator/ (last accessed on 10 July, 2020).

[28] Data on urban population share are obtained from the United Nations, available at: https://population.un.org/wup/ (last accessed on 10 July, 2020).

[29] Per capita GDP is highly correlated with both infant mortality rate ($r = -0.80$) and urban population share ($r = 0.81$).

Main Findings

Table 1 shows the estimated coefficients from models of democracy and militarized conflict initiation. As a reference, the first numerical column shows the results of a univariate model of conflict initiation that ignores the endogeneity of regime type. The second numerical column shows the results of our main trivariate model that jointly estimates conflict initiation and regime types to correct for the endogeneity. By comparing these models, we can see how our inferences change when we correct endogeneity while holding everything else constant.

Let us first look at the results from the "first stage" of our instrumental variables estimation, where challenger's and target's regime type variables are regressed. Consistent with our expectation, the coefficients for the fertility variables are negative and highly statistically significant in both challenger's and target's democracy equations, suggesting that lower fertility rate is associated with democracy. These findings provide strong support for the relevance of the instrument. [30]

[30] Note, however, that variables in the first-stage models can vary only across country and year but not within directed-dyad-years for a given challenger (target) in a given year. In other words, many country-year observations are duplicated in the first stage. This means the standard errors 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 (i.e., dropping any duplicated observations). Table A.5 in the Appendix shows the estimation results. Fertility rate continues to be highly statistically significant ($p < 0.001$) in these country-year models.

Table 1: Models of Democracy and Conflict Initiation, 1960–2010

	(1) Univariate	(2) Trivariate
1: Conflict equation (clog-log)		
Challenger's Democracy	0.133	1.200***
	(0.137)	(0.371)
Target's Democracy	0.357***	0.059
	(0.107)	(0.376)
Joint Democracy	−0.802***	−0.727***
	(0.190)	(0.180)
Military Advantage	0.623***	0.576***
	(0.142)	(0.137)
Challenger's per capita GDP	−0.204***	−0.259***
	(0.069)	(0.067)
Target's per capita GDP	−0.297***	−0.268***
	(0.077)	(0.073)
Challenger's infant mortality	0.157	0.351***
	(0.117)	(0.120)
Target's infant mortality	−0.248*	−0.294**
	(0.127)	(0.125)
Challenger's urbanization	0.011***	0.013***
	(0.004)	(0.004)
Target's urbanization	0.0004	−0.0001
	(0.004)	(0.004)
Peace Years	−0.925***	−0.897***
	(0.036)	(0.041)
Decades fixed-effects	✓	✓
Challenger's region fixed-effects	✓	✓
Target's region fixed-effects	✓	✓
2: Challenger's democracy equation (probit)		
Challenger's Fertility		−0.133***
		(0.042)
Challenger's per capita GDP		0.299***
		(0.065)
Challenger's infant mortality		−0.471***
		(0.105)
Challenger's urbanization		−0.010***
		(0.003)
Decades fixed-effects		✓
Challenger's region fixed-effects		✓
3: Target's democracy equation (probit)		
Target's Fertility		−0.131***
		(0.042)
Target's per capita GDP		0.299***
		(0.065)
Target's infant mortality		−0.474***
		(0.105)
Target's urbanization		−0.010***
		(0.003)
Decades fixed-effects		✓
Target's region fixed-effects		✓
Correlation coefficients		
ρ_{12}		−0.306***
		(0.093)
ρ_{13}		0.071
		(0.090)
ρ_{23}		0.127***
		(0.041)
Number of observations	34,312	34,312

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

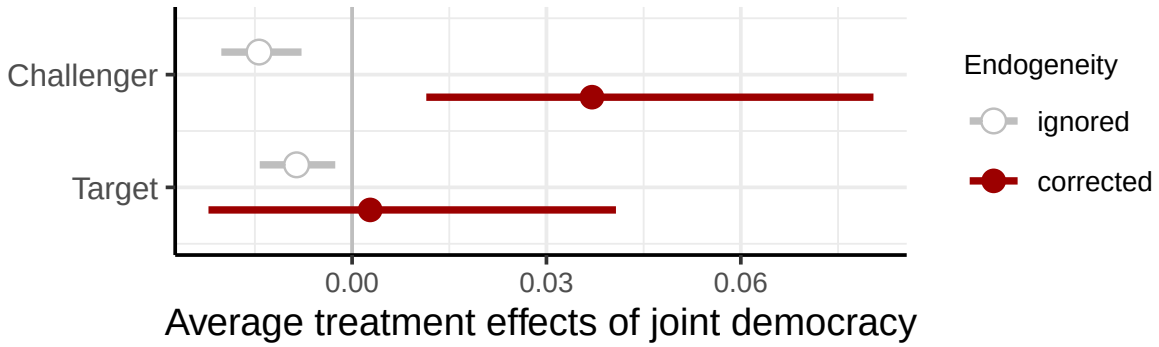

Figure 4: These figures show the estimated average treatment effects of challenger's and target's democracy on militarized conflict. Solid circles shown in red are the point estimates from the trivariate models that correct for the endogeneity problem, whereas hollow circles shown in gray are those from the univariate models that ignore the endogeneity. Horizontal line segments cover the 95% confidence intervals from 1,000 bootstrap replicates.

We now turn to the results from the outcome stage, where militarized conflict initiation is regressed on democracy measures and other covariates. The univariate clog-log model that ignores the endogeneity, shown in column (1) in Table 1, successfully replicates the standard, dyadic democratic peace finding that democracies are peaceful only toward other democracies, as typically reported in previous studies. Note that, while individual democracy measures have either a positive or insignificant coefficient, joint democracy has a negative coefficient that overwhelms the positive coefficients of individual democracy measures in the univariate model. As a result, the univariate model produces a result that, while democracy may increase conflict against a non-democracy, it decreases conflict against a democracy.

To illustrate this, we calculate the average treatment effect of joint democracy for the challenger and for the target based on the univariate model. These effects are calculated by comparing the predicted probabilities of conflict initiation when changing the regime type of self (challenger or target) from non-democracy to democracy, holding constant the regime type of the other (target or challenger) as democracy. [31] Gray, hollow circles

[31] The average effects we present below are averaged over all observations in the estimation sample. The Online Appendix shows the formulae to calculate these effects. We obtain uncertainty estimates using non-parametric bootstrapping.

in Figure 4 show the treatment effects of challenger's and target's democracy. We can see that both of these effects are negative and statistically significant at the 95% confidence level.

But, once we correct the endogeneity, the data no longer support such conclusions. As we see in column (2) in Table 1, negative coefficient for joint democracy no longer overwhelms the positive coefficient of challenger's democracy. This means that challenger's democracy is expected to increase conflict event against a democratic target. Red, solid circles in Figure 4 show the average treatment effects of challenger's and target's democracy, calculated with our trivariate model. We can see that the effect is positive and statistically significant for challenger's democracy, although the effect is indistinguishable from zero for target's democracy.

Whether we correct for endogeneity thus makes a significant difference in our estimates of the effect of joint democracy on conflict. The key to understanding why these changes occur lies in the estimated correlations between the error terms for different equations. The estimated error correlation between equations for conflict and challenger's democracy, ρ_{12}, is negative and statistically significant. This suggests that unobservable or unmeasured determinants of 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 challenger's democracy has a pacifying effect. On the other hand, the estimated error correlation between conflict and target's democracy equations, ρ_{13}, is indistinguishable from zero, suggesting that the endogeneity problem does not seem to operate for target's regime type.

We assess robustness of our results with respect to some of the research design choices we make. First, we estimate models by altering the threshold to code regime type as democracy. We obtain consistent findings when we use polity2 ≥ 5 or polity2 ≥ 7 as the threshold. Second, we estimate the model by dropping per capita GDP and military balance to reduce the danger of post-treatment variable bias. Third, we estimate the model

with smaller data that exclude country-pairs involving China. The Chinese government sought to suppress fertility rate by implementing a "one child" policy. This may have created heterogeneity in the data that could distort the findings. Finally, to assess the generalizability of our findings, we estimate the model with bigger data that add non-contiguous country-pairs that belong to the same geographic regions. We find consistent results from all these additional estimations. [32]

Falsification Tests

Our approach to correct the endogeneity via an IV estimation critically depends on our identification assumptions, the most notable of which is the exclusion restriction — fertility rate must influence militarized conflict exclusively through regime type, conditional on covariates. We have sought to reduce the danger of violating this assumption by controlling for several strong covariates of fertility, but a violation could still occur if there is a causal link from fertility rate to conflict that does not go through regime type. Although we cannot verify this assumption empirically when it is true, we could still falsify it when it is wrong (Labrecque and Swanson 2018). We now assess the plausibility of the exclusion restriction by conducting a series of falsification tests.

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,

[32] All the results are presented in the Online Appendix.

influence international conflict. [33] 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 and second premises, [34] but we doubt that the third premise 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) should lead to democratization. Moreover, premise (iii) assumes that a country's foreign policy *automatically* reflects the preferences 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

[33] 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.

[34] Experimental studies have found somewhat mixed evidence for the first premise. 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.

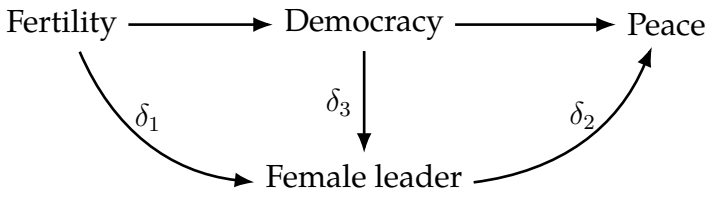

Figure 5: This diagram illustrates a scenario where the exclusion restriction may be violated. The exclusion restriction would be violated when both δ_1 and δ_2 are non-zero at the same time, because that opens up a path that connects fertility and peace that does not go through democracy. The exclusion 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 $\delta_1 = 0$).

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. [35] Figure 5 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., $\delta_2 \neq 0$). [36] 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.

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. [37] The dependent variable is a dummy variable coded as 1

[35] 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.

[36] 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.

[37] We use clog-log link function because having female leaders is a rare event (2.87% of all country-year observations have a female leader).

for a country that has at least one female national leader in a given year, and 0 otherwise. [38]

Table 2: Clog-log Models of Female Leaders

	(3)	(4)	(5)
Fertility rate (δ_1)	-0.042		-0.095
	(0.153)		(0.175)
Democracy (δ_3)		1.166^{***}	1.186^{***}
		(0.325)	(0.318)
Per capita GDP	-0.443^{*}	-0.427^{*}	-0.392
	(0.238)	(0.237)	(0.252)
Infant mortality	-0.615^{**}	-0.638^{**}	-0.542^{*}
	(0.271)	(0.280)	(0.325)
Urbanization	0.010	0.006	0.005
	(0.010)	(0.011)	(0.011)
Decade fixed-effects	✓	✓	✓
Region fixed-effects	✓	✓	✓
Years since last female leader	✓	✓	✓
N	6,595	6,595	6,595

$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$
Estimated coefficients for fixed effects and the intercept are suppressed for brevity.
Standard errors (in parentheses) are clustered at country level.

Table 2 reports the estimation results. Models (3) and (4) include fertility rate and regime type separately, whereas model (5) includes the two variables at the same time. We find no discernible relationship between fertility rate and leader gender (i.e., estimates of δ_1 are indistinguishable from zero), whether or not we control for regime type. On the other hand, democracy seems to make it significant more likely that a country has a female leader (i.e., estimates of $\delta_3 > 0$). 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.

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

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. [39] 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 the findings from our IV model that fertility decreases democracy and that challenger's democracy increases conflict, we expect to find a negative relationship between challenger's fertility and conflict when we conduct a reduced-form analysis on the entire sample.

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 were 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 plausibly 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 emulate such an instance by considering hypothetical principal stratification based on treatment compliance (Angrist, Imbens, and Rubin 1996), imagining our study were an ideal randomized control trial (RCT) on the democratic peace. With regime type being our "treatment" variable (democracy and non-democracy are the possible values) and conflict being the outcome variable in such an imagined RCT, fertility rate would correspond to the "treatment assignment" variable. For example, receiving a low fertility rate would correspond to being assigned to the "treated" group (i.e., assigned to

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

Table 3: Observed and principal strata

	Fertility=Low	Fertility=High
Democracy	Complier / Always-taker	Defier / Always-taker
Non-democracy	Defier / Never-taker	Complier / Never-taker

be democratic). Then, we identify four latent types based on the correspondence between treatment assignment and treatment reception: Compliers are those who would be democratic if and only if fertility rate is low; Defiers are those who would be non-democratic if and only if fertility rate is low; and Always-takers (never-takers) are those who would always (never) be democratic regardless of the assigned fertility level.

Although we cannot directly observe these four ideal types, we could narrow down the possibilities of latent types based on observed treatment and treatment assignment, as shown in Table 3. For example, if we observe a country is democratic and have a low fertility rate, that country is either a complier or an always-taker. [40] Likewise, non-democracies with high fertility rate are either a complier or a never-taker. Given our theoretical argument that lower fertility leads to democracy, it would be reasonable to assume that we do not have a defier in our sample. [41] Then, any observations that fall in the top-right cell or the bottom-left cell in Table 3 would be either an always-taker or a never-taker. Specifically, we identify "always-takers" as democratic country-years with a fertility rate higher than its median value, and "never-takers" as non-democratic country-years with a fertility rate lower than its median value. [42]

[40] Note that it is impossible to narrow this down further purely by observation. To distinguish between compliers and always-takers among those subjects in the top-left cell of the table, we would need to know what the value of the regime type be for the same subject under a counterfactual scenario where the fertility were hypothetically switched to high. The same applies to the other three cells in the table.

[41] This assumption is often called monotonicity assumption. Moreover, given our empirical findings that lower fertility is statistically associated with democracy, there are not many observations in our sample that fall into the top-right cell or the bottom-left cell in the first place.

[42] For example, our classification rule would code China as never-takers for all years after 1975. China has experienced a rapid decline in fertility rate due to the "one-child" policy implemented from 1979 to 2015, while it has been non-democratic during the entire observation period.

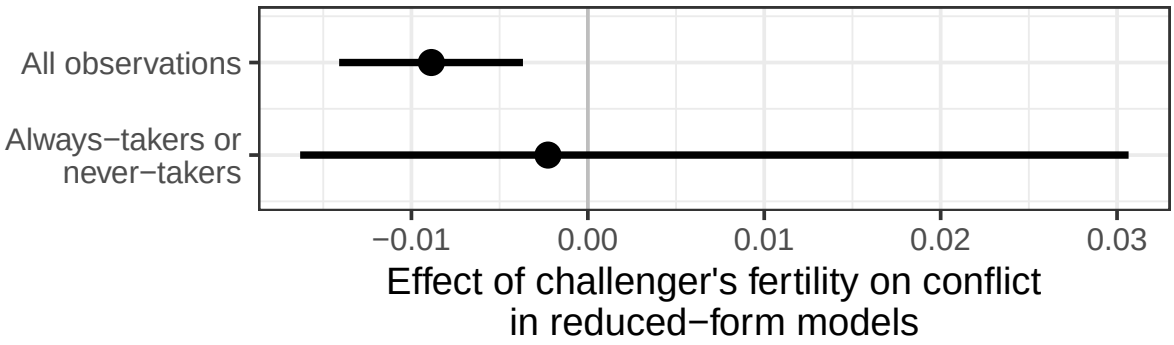

Figure 6: This figure shows the effects of challenger's fertility rate on the probability of militarized conflict estimated with reduced-form models. The top estimate shows the result for all observations, whereas the bottom estimate shows that for a sample that includes only "always-takers" or "never-taker" as defined in the text.

Because the path from fertility to regime type is blocked for always-takers and never-takers, 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. Figure 6 shows the estimated reduced-form relationship between challenger's fertility and conflict for all observations in the sample (top) and for dyads comprising of either always-takers or never-takers (bottom). The x-axis shows the estimated effect of challenger's fertility rate on the probability of militarized conflict initiation in reduced-form models. Effects are calculated by comparing the predicted probabilities of conflict when increasing the value of challenger's fertility rate from its mean value for democracies (2.82) to its mean value for non-democracies (5.17), controlling for target's fertility rate at its mean value for democracies.

As the top estimate shows, the reduced-form model finds a negative relationship between challenger's fertility rate and conflict when estimated for all observations. This is consistent with our findings from the IV model that shows that challenger's democracy has a positive effect on conflict. [43] However, this relationship disappears when we analyze the sample that includes either always-takers or never-takers as defined above. These

[43] The finding is also consistent with the empirical result from Regan and Paskeviciute (2003), who find a negative relationship between fertility rate and conflict. Note that their theoretical argument would expect the opposite relationship.

findings lend support to the exclusion restriction.

Table 4: OLS Models of Fertility

	(6)	(7)	(8)
# of conflicts in the previous year	−0.012		
	(0.025)		
# of conflicts in the previous 3 years		−0.004	
		(0.011)	
# of conflicts in the previous 5 years			−0.003
			(0.007)
Per capita GDP	0.064	0.055	0.043
	(0.092)	(0.095)	(0.095)
Infant mortality	0.794***	0.786***	0.780***
	(0.111)	(0.113)	(0.115)
Urbanization	−0.019***	−0.019***	−0.018***
	(0.004)	(0.004)	(0.004)
Decade fixed-effects	✓	✓	✓
Region fixed-effects	✓	✓	✓
N	6,507	6,266	6,010

$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$
Estimated coefficients for fixed effects and the intercept are suppressed for brevity.
Standard errors (in parentheses) are clustered at country level.

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 fertility rate as the dependent variable and past conflict involvement as an independent variable.

Table 4 shows the results from linear regression models predicting fertility rate for country-year data. In all three models, the dependent variable is fertility rate for a country in a given year. 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 three models,

the coefficients for conflict conflict involvement are indistinguishable from zero, lending support for the exclusion restriction. [44]

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 — on the contrary, democratic countries tend to attack other countries than non-democratic ones do. 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 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? While our findings suggest that it is not the first-cluster (democracy-as-cause) argument, they do not reveal which variant of the second-cluster (reverse causality) or third-cluster (spurious correlation) argument is valid. Nevertheless, we offer an additional criterion to choose among multiple theories of the democratic peace. That is, any theory of the democratic peace must be able to explain not only the observed peace between democracies but also why democratic challenger is more conflict-prone than a non-democratic challenger, even against a democratic target. A promising avenue for future research would be to theorize about these (apparently contradictory) empirical patterns.

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

Another important avenue for future research is to apply the analytical framework of our study to exploring different variables to see if our conclusions generalize beyond our data. 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. It is still possible that a LATE of democracy induced by a different instrument has a negative or an even greater positive effect on conflict — 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., Guido W. Imbens, and Donald B. Rubin. 1996. "Identification of Causal Effects Using Instrumental Variables." *Journal of the American Statistical Association* 91 (434): 444–455.

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.

Braumoeller, Bear F., et al. 2018. "Flexible Causal Inference for Political Science." *Political Analysis* 26 (1): 54–71.

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.

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.

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 Primus Che Chi. 2013. "War and Gender Inequalities in Health: The Impact of Armed Conflict on Fertility and Maternal Mortality." *International Interactions* 39, no. 4 (): 489–510.

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