Fueling Conflict? (De)Escalation and Bilateral Aid

Richard Bluhm* Martin Gassebner† Sarah Langlotz‡

Paul Schaudt§

December 20, 2019

Abstract

This paper studies the effects of bilateral foreign aid on conflict escalation and deescalation. First, we develop a new ordinal measure capturing the two-sided and multifaceted nature of conflict. Second, we propose a dynamic ordered probit estimator that allows for unobserved heterogeneity and corrects for endogeneity. Third, we identify the causal effect of foreign aid on conflict by predicting bilateral aid flows based on electoral outcomes of donor countries which are exogenous to recipients. Receiving bilateral aid raises the chances of escalating from small conflict to armed conflict, but we find little evidence that aid ignites conflict in truly peaceful countries.

Keywords: conflict, foreign aid, political economy, dynamic ordered panel data

JEL Classification: F35, F52, O11, C25

* *Corresponding author*: University of California San Diego and Leibniz University Hannover, e-mail: *bluhm@mak.uni-hannover.de*

†University of Hannover, CESifo, KOF Swiss Economic Institute, e-mail: *gassebner@mak.uni-hannover.de*

‡University of Göttingen, e-mail: *sarah.langlotz@uni-goettingen.de*

§University of St. Gallen, e-mail: *paul.schaudt@unisg.ch*

1. Introduction

Civil conflict is not only one of the main obstacles to development, it also tends to be concentrated in poor countries. About half of all developing countries experienced an armed conflict in which at least 25 people died in a given year over the past four decades—directly or indirectly affecting close to four billion people. At the same time, poor and badly governed states prone to conflict need and receive substantial amounts of development assistance. Bilateral aid averaged about 5% of recipient GDP over the same period, but does this aid appease or fuel conflict?

A large and growing literature examining this question has failed to generate a consensus. Theoretically, the relationship is ambiguous as rising opportunity costs, increasing state capacity, and greater gains from capturing the state are all plausible consequences of development assistance. The empirical evidence is equally divided: several studies find that aid helps, while others maintain that it obstructs peace. Credible evidence is usually limited to specific regions or countries (e.g., the Philippines, Crost et al., 2014), specific types of aid (e.g., U.S. food aid, Nunn and Qian, 2014) or both (e.g., U.S. military aid in Columbia, Dube and Naidu, 2015). Devising a convincing identification strategy for bilateral aid has proven difficult given the well-known limitations of cross-country data.

A notable divide between the theoretical and empirical literature is that the latter pays little attention to the dynamics of conflict. Empirically, conflict is usually considered to be a binary state, although recent theory stresses the importance of smaller conflicts (e.g., Bueno de Mesquita, 2013), different types of violence (e.g., Besley and Persson, 2011b), and conflict cycles (e.g., Rohner et al., 2013, Acemoglu and Wolitzky, 2014). Most papers distinguish between the onset and continuation of conflict, but studying these two transitions separately is an imperfect substitute for analyzing an inherently dynamic problem (Beck et al., 1998). More fundamentally, there is no empirical sense of escalation or deescalation among different conflict intensities when the ordinal nature of conflict is disregarded. Only the case of a switch from peace to conflict and vice versa is usually accounted for. These distinctions matter. As we show in the following, small scale conflicts below the usual minimal threshold of 25 battle-related deaths often start a cycle of violence. In contrast, in our data set, a civil war never broke out in a society that was completely at peace in the year before.

Establishing the causal effect of bilateral aid on the escalation and deescalation of conflict is the key objective of this paper. In essence, we conjecture that neglecting smaller conflicts pollutes the effect of aid on conflict. Civil discontent often first finds its expression in smaller acts of violence with comparatively low opportunity costs. Any violent behavior questions the state's monopoly of violence, satisfying what can be considered the most basic definition of civil conflict. Small conflicts thus act as a signal to the government that some part of society is not content with the current provision, or division, of public goods. In addition, they help potential rebels to get an estimate of how easily they can overcome collective action problems and provide information about the government's repressive capabilities. Foreign aid, in turn, may exacerbate violent tendencies in such environments but not when society is truly at peace.

Our empirical analysis introduces three novelties in order to identify these dynamics. First, we propose a new measure of conflict which captures the gradations of civil violence from peace over intermediate categories to fully fledged civil wars. Second, we develop a dynamic ordered probit framework which allows us to estimate escalation and deescalation probabilities for multiple states. In our approach, the onset, continuation, and the duration of each realization of civil violence are all well defined. We then extend this basic framework to account for unobserved heterogeneity (quasi fixed effects) and correct for the endogeneity of aid (based on Rivers and Vuong, 1988, Wooldridge, 2005, Giles and Murtazashvili, 2013). Third and most importantly, we identify the effect of aid on conflict using characteristics of the electoral system of donor countries. We interact political fractionalization of each donor with the probability of receiving aid to predict bilateral aid flows in a "gravity style" aid equation (Frankel and Romer, 1999, Dreher and Langlotz, 2020). This type of identification strategy is commonly used in the trade and migration literatures but usually relies on structural characteristics of both partner countries. We solely use the variation arising from electoral outcomes in donor countries combined with the likelihood of receiving aid.

Our main results show that the causal effect of foreign aid on the various transition probabilities is heterogeneous and, in some instances, sizable. Foreign aid has a different effect on the probability of experiencing conflict, depending on whether a society was entirely peaceful, already in turmoil, or mired in major civil conflict. Aid does not seem to harm recipient countries by causing conflict across the board. While all estimates suggest that bilateral aid tends to fuel conflict, we find scarce evidence suggesting that foreign aid leads to new eruptions of conflict or that it drives the escalation towards (or the continuation of) civil wars. At face value, the positive signs are also at odds with rising opportunity costs, although it remains difficult to delineate the exact channels.

Our findings suggest that aid can be harmful when given to countries already experiencing violent turmoil just short of the conventional definition of civil conflict. In those cases we find *i)* a strong negative effect on the probability of transitioning back to peace, *ii)* an elevated risk of continued violence, and *iii)* a non-trivial probability of escalating into armed conflict. Donor countries have to be aware of the unintended consequences of giving aid to countries with lingering conflicts.

Our results underscore the importance of carefully modeling the dynamics of conflict. This echoes the recent literature (e.g., Bazzi and Blattman, 2014, Nunn and Qian, 2014, Berman and Couttenier, 2015) but our analysis goes several steps further and generates new insights. Escalation or deescalation, i.e., the switching among different conflict intensities, is a dynamic process and the established binary peace-war typology hides important heterogeneity. What is often coded as peace is not actually peaceful and what influences the decision to fight differs in these situations.

The remainder of the paper is organized as follows. Section 2 discusses the related literature and provides the theoretical background. Section 3 introduces our new ordinal conflict measure. Section 4 outlines our empirical model and identification strategy. Section 5 presents the empirical results and several extensions. Section 6 concludes.

2. Related literature

Civil conflict and foreign aid: The direction of the overall effect of aid boils down to how it changes the calculus of citizens and governments. For citizens, aid may alter the opportunity costs of fighting (e.g., Becker, 1968, Collier and Hoeffler, 2004). On the side of the government, aid may strengthen state capacity (Fearon and Laitin, 2003, Besley and Persson, 2011a) but also increase the value of capturing the state for potential rebels (e.g., Grossman, 1991). Variants of these theories incorporate both channels and try to distinguish between two opposing income effects: having less to fight over but fewer outside options versus fighting over a larger pie but having more to lose. As a result of this heterogeneity, the overall sign of the effect of aid remains theoretically ambiguous.

Most studies in the literature on civil conflict find that aid appeases (e.g., de Ree and Nillesen, 2009, Ahmed and Werker, 2015). Recently, however, evidence to the contrary has been accumulating (e.g., Besley and Persson, 2011b, Nunn and Qian, 2014, Dube and Naidu, 2015). Nunn and Qian (2014), for example, argue that food aid can be used as rebel financing since it can be captured almost instantly. Their results show that U.S. food aid prolongs the duration of conflict but does not predict the onset of conflict. Rising opportunity costs can also lead to an adverse effect of aid. Crost et al. (2014) show that municipalities in the Philippines which are about to receive more aid experience increased rebel activity. Rebels anticipating the impending change in incentives sabotage aid, since successful aid programs reduce support for their cause.

Cycles of violence: The cyclical nature of conflict is receiving increasing attention. Recent theories aim to account for escalation and deescalation cycles in a unified framework. Besley and Persson (2011b) emphasize that one-sided violence by an incumbent aiming to stay in power gives rise to multiple states of violence, ranging from peace over repression to civil war.[1] Rohner et al. (2013) and Acemoglu and Wolitzky (2014) present models where recurring conflicts can happen by accident but are often started when there is a break down of trust or signals are misinterpreted.

[1]Empirically they use ordered outcome models, but do not account for conflict histories, history-dependent effects, or persistence.

Small conflicts matter for a proper understanding of conflict cycles and are often the starting point for further escalation. They are an integral part of rebel tactics for several reasons: *i)* they help to overcome collective action and information problems, *ii)* they have smaller opportunity costs, and *iii)* they can be strategic substitutes to conventional warfare in a long standing rebellion (Bueno de Mesquita, 2013).

A neglect of small conflicts is particularly worrying when it comes to the impact of aid on conflict. The effect of aid may very well be heterogeneous and may depend on the level of violence. This could be the case for at least two reasons. First, aid is not distribution neutral (see, e.g., Dreher et al., 2019, who show that Chinese aid disproportionately flows to the birth region of African leaders). Greater aid flows may increase pre-existing discontent over the allocation of resources. Due to logistical reasons aid is given more often to peaceful regions or regions of low conflict intensity. If aid is primarily targeted at such regions, resentment may fortify in unprivileged areas, where violence persists. Opportunity costs erode and rebels controlling such a region may be able to recruit others more easily. Second, if a country is entirely peaceful, the government is less likely to divert development aid or freed-up funds to the military. If there is a lingering conflict, on the other hand, the incumbent government might continue to invest in the military to repress or discourage rebellion (Besley and Persson, 2011a). Hence, the effect of aid on state capacity differs depending on the level of violence.

Causal identification: The simultaneity of aid and conflict makes causal identification notoriously difficult. The strong correlation of low GDP per capita and civil strife is one of the most robust findings in the literature (e.g., Fearon and Laitin, 2003, Blattman and Miguel, 2010). Underdevelopment—with all that it entails—is the *raison d'être* of development aid (see, e.g., Brückner, 2013, on the simultaneity of these two). As a result, the effect of aid is likely to be biased upwards if aid is primarily given to countries in need, or biased downwards if donors are driven by political motives or reduce aid in light of the logistical challenges created by conflict. Biases could also result from third factors influencing aid and conflict simultaneously, such as political and economic crises, or (systematic) measurement errors.

To address this endogeneity Nunn and Qian (2014) use lags of U.S. wheat production interacted with each recipient's frequency of receiving aid as an instrument for U.S. food aid. We extend the spirit of their identification strategy to all major bilateral donors, with the explicit aim of drawing conclusions that go beyond the (limited) effects of food aid given by one large donor. Much of the ground work has been done in Dreher and Langlotz (2020) who first introduce political fractionalization interacted with the probability of receiving aid as an instrument for bilateral aid flows in the context of growth regressions. We describe this strategy in more detail below.

3. Data

We study the occurrence of civil violence in 125 developing countries over the period from 1975 to 2010. A list of the included countries and summary statistics of all variables can be found in Section A of the Online Appendix.

An ordinal measure of conflict: A distinct feature of the civil conflict literature is its crude measurement of conflict. The industry standard is to first count the number of battle-related deaths (BDs) and then to create dummy variables indicating the surpassing of one of two thresholds (25 or 1,000 BDs) for the first time (conflict onset) or for any given year other than the first (continuation or ending). Clearly, a key concern motivating this choice is noise in the underlying raw data and theoretical ambiguity about what constitutes "conflict."

We propose a new ordinal measure of conflict with four states. For comparability, we begin with the standard UCDP-PRIO measure of civil conflict (Gleditsch et al., 2002). UCDP-PRIO defines civil conflict as a contested incompatibility that concerns the government or a territory in which armed force between two parties, one of which is the government, results in at least 25 BDs per annum. We call conflicts that reach this state but do not exceed 1,000 BDs in a given year '*armed conflict*.' At the top, we add a category called '*civil war*' if there are more than 1,000 BDs. At the bottom, we create the category '*small conflict*' using data from the Cross-National Time-Series

Data Archive (CNTS) on government purges, assassinations, riots and guerrilla warfare (Banks and Wilson, 2015).[2] All of these categories are manifestations of civil conflict, albeit on a lower intensity level. The UCDP-PRIO data are generally considered to be more rigorously constructed than the CTNS data but do not report low intensity events. We solely use the CNTS data to identify if there has been any violence below the 25 BDs threshold of armed conflict. We include observations of the CNTS data that are comparable to the type of conflict we consider in the above categories, i.e., conflicts between two parties one being the state (two-sided, state-centered).[3] Only a society without any event is assigned to the category '*peace*.' As a whole, the countries in our sample spend about one third of all years in conflict at various intensities and about two thirds of all years in peace (see Figure I).

A key advantage of our approach is that the number of armed conflicts and civil wars in our sample are identical to the UCDP-PRIO measure. Hence, our results are comparable with existing studies and differ mainly due to the definition of peace. We distinguish between truly peaceful observations and those with irregular violence below the conventional thresholds. This conservative approach of changing existing measures implies that our ordinal measure can be easily compared to the industry standard and its interpretation is straightforward. We avoid weighting procedures such as those used by the composite index of the CNTS data set. We also deliberately refrain from mixing flow and stock variables to measure different conflict intensities, such as taking the cumulative amount of BDs to create intermediate levels of armed civil conflict (e.g., Esteban et al., 2012, Bazzi and Blattman, 2014). Measures including both flow and stock variables do not allow us to study escalation and deescalation since they have absorbing terminal states. The Online Appendix presents the case of the Sri Lankan civil war to illustrate

[2] The precise definitions of our variables from the Databanks User's Manual are as follows. Purges: Any systematic elimination by jailing or execution of political opposition within the ranks of the regime or the opposition. Assassinations: Any politically motivated murder or attempted murder of a high government official or politician. Riots: Any violent demonstration or clash of more than 100 citizens involving the use of physical force. Guerrilla Warfare: Any armed activity, sabotage, or bombings carried on by independent bands of citizens or irregular forces and aimed at the overthrow of the present regime.

[3] In the case of riots this may not be obvious from the variable definition, but the large riots recorded in the CNTS data usually involve violent clashes between anti-government protesters with (pro-)government forces. They are what incumbents react to with repression. For a prototypical example, see Yemen in 2011 (http://www.nytimes.com/2011/02/15/world/middleeast/15yemen.html).

the benefits of our coding in more detail.

Table I shows the unconditional transition probabilities as they are observed in our data. This simple exercise already allows us to make three worthwhile points. First, the cyclical nature of conflicts is clearly visible but there is not a single country in our data set where peace immediately preceded civil war. Second, our coding of small conflict achieves a credible and important separation of the lower category. Peace is now very persistent and, if anything, a transition to a small conflict is most likely. Small conflict is a fragile state which often reverts back to peace, is not particularly persistent, but does sometimes erupt into more violent states. Third, higher intensity conflicts are once again more persistent. These observations match up well with the literature, in particular, the use of irregular means to increase mobilization for a future conventional campaign and increased persistence as outside opportunities erode (Bueno de Mesquita, 2013).

Bilateral aid flows and controls: Our main independent variable is official development aid (ODA) disbursed by 28 bilateral donors of the OECD Development Assistance Committee (DAC). ODA refers to flows that are *i)* provided by official agencies to developing countries and multilateral institutions, *ii)* have economic development and welfare as their main objective, and *iii)* have a concessional character. The last condition reflects that the grant element should be at least 25%.[4] We use net ODA flows which include loan repayments since these reduce the available funds. In the Online Appendix, we also consider multilateral aid. All flows are normalized by GDP.

The data for government and legislative fractionalization (in donor countries) are from Beck et al. (2001). For the set of core controls, we include the log of population (from the World Development Indicators) to capture the scale effect inherent in conflict incidence and the log of GDP (from the Penn World Table 7.1).

[4]Other official flows (OOF) includes flows by the official sector with a grant element of less than 25% or flows that are not primarily aimed at development. Table D-4 in the Online Appendix presents results based on these flows.

4. Empirical strategy

A. Conflict histories

Dynamic switches among multiple states cannot be meaningfully estimated with linear models. Beck et al. (1998) show that separately specifying models of onset and ending of war is equivalent to a dynamic model of war incidence. However, many more linear models would be needed to study the transition among multiple states. The result would be unstable parameter estimates that are inefficiently estimated, potentially biased, and difficult to interpret. Further, if we believe that there is an underlying latent variable ('conflict') which is observed as an ordered outcome, then separate regressions can violate known parameter restrictions.[5] Hence, a non-linear framework is needed.

Some notation is in order to help fix ideas. As typical in an ordered setting, we observe a conflict outcome c_{it} which takes on $J + 1$ different values in country i at time t. A specific outcome is $j \in \{0, 1, \ldots, J\}$. The outcomes are ordered by intensity (i.e., peace, small conflict, armed conflict, civil war) and are generated by a continuous latent variable c^*_{it} with J cut points $\alpha_1 < \cdots < \alpha_j < \cdots < \alpha_J$ to be estimated later. The first outcome is $c_{it} = 0$ if $-\infty < c^*_{it} < \alpha_1$, the intermediate outcomes are $c_{it} = j$ if $\alpha_j < c^*_{it} < \alpha_{j+1}$ with $0 < j < J$, and the last outcome is $c_{it} = J$ if $\alpha_J < c^*_{it} < \infty$.

Next, define the associated $J \times 1$ vector of one period conflict histories as $\mathbf{h}_{i,t-1} \equiv (h_{1,i,t-1}, \ldots, h_{j,i,t-1}, \ldots, h_{J,i,t-1})'$. The typical element of $\mathbf{h}_{i,t-1}$ is $h_{j,i,t-1} \equiv 1[c_{i,t-1} = j]$, that is, an indicator of whether the past outcome is identical to outcome j.

Contrary to the standard approach, our latent variable model of interest has a full set of history dependent effects

$$c^*_{it} = \mathbf{x}'_{it}\boldsymbol{\beta} + \mathbf{h}'_{i,t-1}\boldsymbol{\rho} + (\mathbf{x}_{it} \otimes \mathbf{h}_{i,t-1})'\boldsymbol{\gamma} + \mu_i + \varepsilon_{it} \tag{1}$$

where $\mathbf{x}_{it}$ is a column vector of regressors without a constant, $\mathbf{h}_{i,t-1}$ is defined above, and

[5] This is a version of the misnamed "parallel regression assumption" in ordered probit models. If the outcome is an ordered response, then the predicted probabilities of falling below a certain cut point *must* be increasing in the outcome j for all values of the covariates (Wooldridge, 2010, p. 658). If all the coefficients can vary in each state, then this meaningless result cannot be ruled out.

the Kronecker product simply accounts for all possible interactions between $\mathbf{x}_{it}$ and $\mathbf{h}_{i,t-1}$. We include country level unobserved effects, μ_i, whose identification we discuss below. Typically we will partition the vector $\mathbf{x}_{it} = (\mathbf{x_1}'_{it}, \mathbf{x_2}'_{it})'$, so that some variables are history dependent and others are not (e.g., proxy controls and time dummies). We are only interested in the estimated coefficients inasfar as they define the relevant probabilities.

Conditional on the covariates and the conflict history we have three different types of outcome probabilities: $\Pr[c_{it} = 0|\mathbf{x}_{it}, \mathbf{h}_{i,t-1}] = \Pr[c^*_{it} \leq \alpha_1|\mathbf{x}_{it}, \mathbf{h}_{i,t-1}]$, $\Pr[c_{it} = j|\mathbf{x}_{it}, \mathbf{h}_{i,t-1}]$ $= \Pr[\alpha_j < c^*_{it} \leq \alpha_{j+1}|\mathbf{x}_{it}, \mathbf{h}_{i,t-1}]$, and $\Pr[c_{it} = J|\mathbf{x}_{it}, \mathbf{h}_{i,t-1}] = \Pr[c^*_{it} > \alpha_J|\mathbf{x}_{it}, \mathbf{h}_{i,t-1}]$. We have to be more explicit in the notation since we are interested in the transition and continuation probabilities of the various states. For simplicity, just focus on the j-th intermediate outcome where $0 < j < J-1$, then w.l.o.g. we can define continuation, escalation and deescalation from an initial state $j+p$ to outcome j as:

$$\begin{aligned}\Pr[c_{it} = j|\mathbf{x}_{it}, h_{j+p,i,t-1} = 1] =& F\left[\alpha_{j+1} - \mathbf{x}'_{it}\boldsymbol{\beta} - \rho_{j+p} - (\mathbf{x}_{it} \times h_{j+p,i,t-1})'\boldsymbol{\gamma}_{j+p} - \mu_i\right] \\ &- F\left[\alpha_j - \mathbf{x}'_{it}\boldsymbol{\beta} - \rho_{j+p} - (\mathbf{x}_{it} \times h_{j+p,i,t-1})'\boldsymbol{\gamma}_{j+p} - \mu_i\right]\end{aligned} \tag{2}$$

where we have escalation if $p < 0$, continuation if $p = 0$ and deescalation if $p > 0$. The case of $p = 0$ is often also called 'persistence.' $F(\cdot)$ is some continuous symmetric c.d.f. which is defined by the distribution of the error terms, ε_{it}.

The purpose of this entire exercise is to be able to define the partial effect of a particular $x_{k,it} \in \mathbf{x}_{it}$ on one of the transition probabilities defined above. It should now be straightforward to see that these are the derivatives of a particular probability with respect to $x_{k,it}$. For example, in the case of continuing in the past state j we have

$$\begin{aligned}\frac{\partial}{\partial x_k}\left(\Pr[c_{it} = j|\mathbf{x}_{it}, h_{j,i,t-1} = 1]\right) =& (\beta_k + \gamma_{j,k})\Big(f\left[\alpha_j - \mathbf{x}'_{it}\boldsymbol{\beta} - \rho_j - (\mathbf{x}_{it} \times h_{j,i,t-1})'\boldsymbol{\gamma}_j - \mu_i\right] \\ &- f\left[\alpha_{j+1} - \mathbf{x}'_{it}\boldsymbol{\beta} - \rho_j - (\mathbf{x}_{it} \times h_{j,i,t-1})'\boldsymbol{\gamma}_j - \mu_i\right]\Big),\end{aligned} \tag{3}$$

where $f(\cdot)$ is the p.d.f. of $F(\cdot)$.

We still lack a formal definition of state dependence. In binary models, state dependence is the probability of an event happening when the event happened before minus the probability of the event when it did not happen before net of all other observed and unobserved factors. With ordered outcomes it is no longer that simple. We need to account for the fact that there are several ways of entering into a particular state. Following Cappellari and Jenkins (2004), we estimate state dependence as the difference between experiencing a particular state if it has occurred before and a weighted average of the ways of entering this state when it has not occurred before.

Formally, define state dependence in state j as follows:

$$S_j = (NT)^{-1} \sum_{i}^{N} \sum_{t}^{T} \left(\Pr[c_{it} = j | \mathbf{x}_{it}, h_{j,i,t-1} = 1] - \sum_{r \neq j} \omega_{rj} \Pr[c_{it} = j | \mathbf{x}_{it}, h_{r,i,t-1} = 1] \right), \tag{4}$$

where the weights, ω_{rj}, are the normalized class frequencies (the number of observations that can potentially make the switch, normalized to sum to unity). We expect state dependence to increase with higher conflict intensities. The higher the level of conflict, the more difficult it becomes to leave states that have a destructive nature, whereas peace should also be self-reinforcing.

B. Dynamic ordered probit with endogeneity

Identification of endogenous regressors and their partial effects under the presence of heterogeneity and first-order dynamics is not trivial in non-linear settings. Researchers often opt for linear instrumental variable methods to keep things simple, but here we trade simplicity for a better understanding of the dynamics.

To model the ordered conflict outcome, we combine correlated random effects (CRE) and a control function (CF) approach with dynamic panel ordered probit models. Dynamic models with correlated random effects where all regressors are strictly exogenous have been studied by Wooldridge (2005), among others, and endogeneity was introduced into these types of dynamic binary choice models by Giles and Murtazashvili (2013). To the best of our knowledge, we are the first to employ a CRE approach with an endogenous regressor in an dynamic ordered setting. A noteworthy limitation of this approach is that it does not work with unbalanced panels.

We incorporate two specific features into the general formulation considered above. First, we add an endogenous regressor (the ratio of bilateral aid to GDP) and, second, we interact this variable with the one period conflict history. We do not consider other interactions. Hence, our model of interest becomes

$$c^*_{1it} = \mathbf{z}'_{1it}\boldsymbol{\beta}_1 + \beta_2 a_{2it} + \mathbf{h}'_{1i,t-1}\boldsymbol{\rho} + (a_{2it} \times \mathbf{h}_{1i,t-1})'\boldsymbol{\gamma} + \mu_{1i} + \lambda_{1t} + u_{1it} \tag{5}$$

where $\mathbf{z}_{1it}$ is a column vector of strictly exogenous variables, a_{2it} is the endogenous aid to GDP ratio, λ_{1t} are time dummies, and everything else is defined as before. We added subscripts to each variable or vector if they belong to the main equation of interest (1) or the reduced form (2). We assume that the model is dynamically complete once the first-order dynamics are accounted for and that the error term is free of serial correlation. We relax this assumption later when we discuss extensions to the model. The process starts at $s < 0$ and is observed over $t = 0, \ldots, T$. We always lose the first period, so in eq. 5 and from now on estimation runs over $t = 1, \ldots, T$.

The endogenous aid to GDP ratio has the following linear reduced form

$$a_{2it} = \mathbf{z}'_{1it}\boldsymbol{\alpha}_1 + \mathbf{z}'_{2it}\boldsymbol{\alpha}_2 + \mu_{2i} + \lambda_{2t} + u_{2it} \tag{6}$$

where $\mathbf{z}_{2it}$ is a vector of instruments that is relevant and excluded from the main equation. Our instrument is generated from bilateral regressions. We discuss its construction in detail in the next subsection.

We assume that the reduced form heterogeneity can be expressed as $\mu_{2i} = \bar{\mathbf{z}}'_i\boldsymbol{\psi} + b_{2i}$, where $b_{2i}|\mathbf{z}_i \sim \mathcal{N}(0, \sigma^2_{b_2})$ and $\mathbf{z}_i \equiv (\mathbf{z}'_{1it}, \mathbf{z}'_{2it})' \equiv (\mathbf{z}'_{i1}, \mathbf{z}'_{i2}, \ldots, \mathbf{z}'_{iT})'$ is a vector of all strictly exogenous variables in all time periods. Plugging this into eq. 6 gives

$$a_{2it} = \mathbf{z}'_{1it}\boldsymbol{\alpha}_1 + \mathbf{z}'_{2it}\boldsymbol{\alpha}_2 + \bar{\mathbf{z}}'_i\boldsymbol{\psi} + \lambda_{2t} + \nu_{2it} \tag{7}$$

where $\nu_{2it} = b_{2i} + u_{2it}$ is the new composite error term. It is well known that the coefficients on the time-varying covariates in eq. 7 are numerically equivalent to the linear fixed effects model, making this a very robust specification (Wooldridge, 2010, p. 332).

Following Rivers and Vuong (1988) and Giles and Murtazashvili (2013), joint normality of (u_{1it}, u_{2it}) conditional on $\mathbf{z}_i$ with $Var(u_{1it}) = 1$, $Cov(u_{1it}, u_{2it}) = \tau$, and $Var(u_{2it}) = \sigma^2_{u_2}$ implies that we can rewrite our model of interest as

$$c^*_{1it} = \mathbf{z}'_{1it}\boldsymbol{\beta}_1 + \beta_2 a_{2it} + \mathbf{h}'_{1i,t-1}\boldsymbol{\rho} + (a_{2it} \times \mathbf{h}_{1i,t-1})'\boldsymbol{\gamma} + \mu_{1i} + \lambda_{1t} + \omega u_{2it} + \varepsilon_{1it}, \tag{8}$$

where we define $\omega = \tau/\sigma_{u_2}$.

Note that $u_{1it} = \omega u_{2it} + \varepsilon_{1it} = \omega(\nu_{2it} - b_{2i}) + \varepsilon_{1it}$, so our equation of interest is contaminated by both the first stage errors and the associated unobserved heterogeneity. The role of ν_{2it} is to "correct" for the contemporaneous endogeneity between the two equations, while b_{2i} allows for feedback from the unobserved effect in the reduced form.

Let $b_{1i} = \mu_{1i} - \omega(\nu_{2it} - u_{2it})$ be the composite unobserved effect and note that this does not depend on t because $\nu_{2it} = b_{2i} + u_{2it}$ by definition. Then the key question in non-linear dynamic models is what assumptions do we make about how the composite heterogeneity relates to the initial conditions $\mathbf{h}_{i0}$, the covariates $\mathbf{z}_i$ and the reduced form errors in all periods $\boldsymbol{\nu}_{2i}$?

Assuming that the heterogeneity only relates to the reduced form errors gives rise to a random effects specification with Mundlak terms for the first stage residuals. Assuming that the composite heterogeneity is a linear function of all three gives rise to a dynamic correlated random effects approach. The initial conditions are not ignorable and have repercussion towards how flexibly we must treat the unobserved heterogeneity (Wooldridge, 2005). In our case, some unobserved covariate of conflict is likely to be correlated with whether a country was initially in conflict or peace, so that independence of $\mathbf{h}_{i0}$ and $\mathbf{z}_i$ is unlikely.

Following Giles and Murtazashvili (2013), we assume that $b_{1i}|\mathbf{z}_i, \mathbf{h}_{i0}, \boldsymbol{\nu}_{2i} \sim \mathcal{N}(\mathbf{z}'_i\boldsymbol{\delta}_0 + \mathbf{h}'_{i0}\boldsymbol{\delta}_1 + \boldsymbol{\nu}'_{2i}\boldsymbol{\delta}_3, \sigma^2_d)$. This homoskedastic normal distribution implies that the composite heterogeneity is a linear function: $b_{1i} = \mathbf{z}'_i\boldsymbol{\delta}_0 + \mathbf{h}'_{i0}\boldsymbol{\delta}_1 + \boldsymbol{\nu}'_{2i}\boldsymbol{\delta}_3 + d_{1i}$ where $d_{1i}|\mathbf{z}_i, \mathbf{h}_{i0}, \boldsymbol{\nu}_{2i} \sim$

$\mathcal{N}(0, \sigma_d^2)$. Plugging this into eq. 8 gives the final equation

$$\begin{aligned} c_{1it}^* = \; & \mathbf{z}_{1it}'\boldsymbol{\beta}_1 + \beta_2 a_{2it} + \mathbf{h}_{1i,t-1}'\boldsymbol{\rho} + (a_{2it} \times \mathbf{h}_{1i,t-1})'\boldsymbol{\gamma} + \omega\nu_{2it} \\ & + \lambda_{1t} + \mathbf{z}_i'\boldsymbol{\delta}_0 + \mathbf{h}_{i0}'\boldsymbol{\delta}_1 + \boldsymbol{\nu}_{2i}'\boldsymbol{\delta}_3 + d_{1i} + \varepsilon_{1it}, \end{aligned} \tag{9}$$

which can be estimated by standard random effects ordered probit along with the cut points α_j which will result in scaled parameters (e.g., $\boldsymbol{\beta}_1/\sqrt{(1+\sigma_{d_1}^2)}$ and so on, assuming the usual normalization of $Var(\varepsilon_{1it}) = 1$ is applied).

If the instrument were observed as is, then this approach would imply a two-step estimator. Here we implement this estimator in three steps: *i)* we first generate the instrument from bilateral regressions as described below, *ii)* we then estimate the reduced form given in eq. 7, obtain an estimate of the residuals ($\hat{\nu}_{2it}$) and the reduced form errors in all periods ($\hat{\boldsymbol{\nu}}_{2i}$), and then *iii)* plug these into eq. 9. The standard errors are bootstrapped over all three steps to account for the generation of the instrument and the subsequent estimation of the residuals in the first step. Note that the CF approach does not require interactions with the residuals unlike IV methods, making it somewhat less robust but potentially much more efficient (Wooldridge, 2010, p. 128).

In our case T is moderately large which has two implications. First, adding a new time-varying control variable means adding T additional regressors. Second, the initial conditions problem is not likely to be severe. Rabe-Hesketh and Skrondal (2013) provide simulation results for different ways of specifying the conditional density of the unobserved effect in the dynamic binary probit model. Inspired by their study, we experimented with constraints that can be placed on the two sequences $\mathbf{z}_i$ and $\hat{\boldsymbol{\nu}}_{2i}$. Our results suggest that allowing only the first few periods to have an independent effect and constraining the rest to the time averages yields results that are almost indistinguishable from the full model.[6]

The average partial effects (APEs) are derivatives of the expectation of our specification with respect to the distribution of b_{1i}. The APEs can be different for each

[6] We conserve degrees of freedom by splitting the two vectors, so that in the case of the exogenous variables we have $\mathbf{z}_i^+ = (\mathbf{z}_{i1}', \mathbf{z}_{i2}', \ldots, \mathbf{z}_{iR}', \bar{\mathbf{z}}_i^{+'})'$ where $R < T$ and $\bar{\mathbf{z}}_i^+ = \frac{1}{T-R-1}\sum_{t=R+1}^{T} \mathbf{z}_{it}$ is the time average after period R. The residual sequence, $\boldsymbol{\nu}_{2i}^+$, is computed analogously. Our results are not sensitive to the choice of R, as long as the first period is allowed to have its own coefficients. We typically set $R = 4$. We also included $\mathbf{z}_{i0}$ to little effect (as suggested by Rabe-Hesketh and Skrondal, 2013).

t. We usually average across all observations to obtain a single estimate.

C. Identification

We use political fractionalization in donor countries interacted with the probability of receiving aid as our primary source of exogenous variation at the donor-recipient level. Dreher and Langlotz (2020) show that government fractionalization interacted with this probability is a strong instrument for bilateral aid. Government fractionalization is defined as the probability that any two randomly-chosen deputies of the parties forming the government represent different parties (Beck et al., 2001). Fractionalized governments require coalitions to govern, which leads to higher expenditures through logrolling (favors) in the budgeting process, including higher foreign aid budgets and subsequent disbursements (Dreher and Fuchs, 2011, Dreher and Langlotz, 2020).

Most studies analyzing the effects of political fractionalization on government spending focus on parliamentary systems with proportional representation. This is because coalition governments are more likely to be generated by some systems rather than others. Electoral rules, in particular first-past-the-post (FPTP) rules, define if government can be fractionalized at all or if there is a single-party government which negotiates the budget process in some form of reconciliation process with the legislative body. Persson et al. (2007) present a model along these lines where majoritarian elections usually lead to single party government and less spending in equilibrium than proportional elections. Hence, we prefer government fractionalization over fractionalization of the legislature as an instrument in parliamentary systems with proportional representation.[7] For the few donors with FPTP systems—i.e., Canada, the UK, and the U.S.—we use legislative fractionalization as our preferred source of exogenous variation.[8]

The interaction with the probability of receiving aid then introduces variation across recipients. An interaction of this endogenous probability with an exogenous variable is

[7] Legislative fractionalization is defined similarly to government fractionalization. It gives the probability of randomly picking two deputies from the legislature that belong to different parties.

[8] France is an interesting case as it is a mixed system with two-round runoff voting. However, both government and legislative fractionalization vary for France. In a robustness test we also treat France in the same way as Canada, the UK, and the U.S. without a material impact on the results.

itself exogenous, provided we include country and time fixed effects. Just as in Nunn and Qian (2014), our identification strategy can be related to a difference-in-difference approach. We essentially compare the effects of aid induced by changes in political fractionalization in donor countries among regular and irregular aid recipients. We later also examine the parallel trends assumption inherent in our approach.

Applying this in a bilateral setting requires aggregating the bilateral variation in the instruments to the recipient-year level. We opt for a regression approach in which we predict aid bilaterally from the best linear combination of the two interacted instruments and then aggregate the bilateral predictions. Specifically, we predict aid from donor j to recipient i in year t using the following regression:

$$a_{3ijt} = \theta_0 g_{3jt} + \theta_1 (g_{3jt} \times \bar{p}_{3ij}) + \xi_0 l_{3jt} + \xi_1 (l_{3jt} \times \bar{p}_{3ij}) + \mu_{3ij} + \lambda_{3t} + \epsilon_{3ijt} \tag{10}$$

where g_{3jt} is government fractionalization, l_{3jt} legislative fractionalization and $\bar{p}_{3ij}$ is the pairwise probability of receiving aid. As discussed above g_{3jt} is typically zero in FPTP systems. We set all FPTP observations of $g_{3jt} = 0$ and $l_{3jt} = 0$ in non-FPTP systems, in order to only utilize the relevant political type of political heterogeneity. The time-invariant probability is defined as $\bar{p}_{3ij} = \frac{1}{T} \sum_t^T \mathbf{1}[a_{3ijt} > 0]$, so that it contains the fraction of years in which recipient i received a positive amount of aid from donor j. We again added subscripts to indicate that this equation (3) precedes the others with index (2) and (1). We do not need to control for the endogenous level of $\bar{p}_{3ij}$ as it is captured by the recipient-donor fixed effects, μ_{3ij}. We then aggregate the predicted bilateral aid from eq. 10 across all donors in order to get predicted aid as a share of GDP at the recipient-year level. Hence, $\hat{a}_{2it} = \sum_j \hat{a}_{3ijt}$ is the instrument in eq. 7.

We may worry about what variation actually ends up in our constructed instrument. To be clear, it consists of three different components: *i)* the estimated donor-recipient fixed effects aggregated over all donors, or $\sum_j \hat{\mu}_{3ij}$, *ii)* the estimated effects of those donor characteristics that do not vary across recipients and the time dummies aggregated over all donors, or $\sum_j \hat{\theta}_0 g_{3jt} + \sum_j \hat{\xi}_0 l_{3jt} + J\hat{\lambda}_{3t}$, and, finally, *iii)* the exogenous variation introduced by the two interaction terms aggregated over all donors, or $\sum_j \hat{\theta}_1 (g_{3jt} \times \bar{p}_{3ij}) + \sum_j \hat{\xi}_1 (l_{3jt} \times$

$\bar{p}_{3ij}$). The first two are potentially endogenous, but we control for their influence in the estimation that follows. Donor fractionalization is the same across all recipients and will be swept out by the fixed effects (or time averages) in the reduced form equation. Similarly, everything but the interaction terms will be swept out by the recipient effects and time effects.

5. Results

A. Bilateral estimation

We begin by briefly discussing the bilateral regression which we use to construct the instrument. Recall that we regress aid received by each recipient from a particular donor on political fractionalization, its interaction with the probability of receiving aid, and a full set of country and time fixed effects. We estimate these models with the fraction of aid in GDP as the dependent variable (not in logs, since negative flows occur when loan repayments exceed new inflows).[9]

The regression is estimated over 4,116 bilateral donor-recipient relations for which we have data, yielding a total of 129,348 observations.[10] These results are not causal. They only serve to "translate" the exogenous variation in donor characteristics into changes in aid disbursements at the recipient level, depending on how strongly a particular recipient depends on aid from a particular donor.

The estimated coefficients of our variables of interest are as follows (standard errors accounting for bilateral correlation at the donor-recipient level are reported in parentheses below the coefficients):

$$\hat{a}_{3ijt} = \cdots - \underset{(0.014)}{0.043}\, g_{3jt} + \underset{(0.058)}{0.227}\, (g_{3jt} \times \bar{p}_{3ij}) + \underset{(1.407)}{2.564}\, l_{3jt} - \underset{(1.426)}{2.936}\, (l_{3jt} \times \bar{p}_{3ij}). \qquad (11)$$

[9] Negative flows also prevent us from using Poisson Pseudo Maximum Likelihood models commonly used for estimating gravity equations.

[10] We do not constrain this estimation to the balanced sample to arrive at the best possible estimate of this relationship.

The coefficients on the interaction terms are highly significant.[11] Note that the negative sign on the second interaction coefficient is misleading. In both cases, increasing political fractionalization leads to more aid disbursements for nearly all of the sample. There is interesting heterogeneity across political systems. Fractionalized parliamentary systems give more aid to regular recipients, whereas divided majoritarian systems give more aid to irregular recipients.

The effects of political fractionalization are not as large as a cursory glance at the coefficients may suggest. To see this, consider a 10 percentage points increase of political fractionalization in a donor country when a recipient receives aid about two thirds of the time. Eq. 11 predicts that this increases the aid to GDP ratio by about 0.01 percentage point for aid from proportional systems ($0.1 \times [-0.043 + 0.227 \times 2/3] \approx 0.01$) and about 0.06 percentage points for aid from majoritarian systems ($0.1 \times [2.564 - 2.936 \times 2/3] \approx 0.06$). The increase in majoritarian systems tends to be larger, in part because it is estimated based solely on three of the biggest donors. The cluster-robust F-statistic of the interaction terms is about 10.83. Note that the constructed instrument will turn out to be considerably stronger once we aggregate to the recipient (country) level. We essentially add up many small changes in the aid to GDP ratio in any given year.[12]

B. Reduced form of aid

We now turn to country level estimates of the first stage relationship. Table II shows three reduced form regressions for aid to GDP which we obtain by estimating the equivalent fixed effects model of eq. 7. The residuals from these models are used as control functions in the main specifications which we estimate further below. The sample is now balanced at $T = 36$ (minus the initial period) and $N = 125$. This constitutes a much larger sample relative to the typical study in this field which often focuses exclusively on Sub-Saharan Africa or loses observations due to the inclusion of many controls. Our data contains countries experiencing some of the most severe and longest-running civil conflicts (e.g.,

[11] Clustering at the recipient level yields very similar results. Since we only have 28 donors, the standard errors widen noticeably when we cluster at the donor level or donor and recipient level.

[12] We repeated this estimation using net aid including other official flows. The results are qualitatively and statistically similar (not reported, available on request).

Afghanistan, Iraq, Pakistan and many more). Generated regressors deliver consistent estimates but inference depends on the uncertainty introduced in the bilateral stage, which is why we compute bootstrap standard errors (i.e., by sampling over recipients in the bilateral stage, generating the instrument in each iteration, aggregating to a country-year panel, and then estimating the first stage).

Two things stand out in Table II. First, the estimated coefficients on the instruments in all columns are always larger than one. Depending on the specification, a one percentage point increase in the predicted aid to GDP ratio leads to about a 1.3 percentage point increase in actual aid to GDP. Adding other controls moves the estimated coefficients a bit closer to unity. Including heterogeneous linear trends for each initial conflict state changes the results only marginally. The size of the coefficient has a straightforward interpretation. If the constructed instrument overpredicts the quantity in question, then the coefficient will be below unity, and *vice versa*. Not surprisingly, our aggregation of predicted bilateral flows tends to undershoot actual aid to GDP ratios and therefore has a multiplier above unity. Second, the aggregated instrument is highly relevant. The F-statistics based on the bootstrapped covariance matrix always exceed the conventional level of about ten by an order of magnitude (which is also not unusual in comparable applications).[13] Our instrument does not draw its power from any one donor or recipient,[14] or settings where many donors are active at the same time.[15]

A number of other concerns could be raised regarding the strength and validity of our identification strategy. Fractionalized governments and legislatures could be giving more aid to countries that are politically closer, more open to trade or that receive a lot of foreign direct investment. Any (conditional) correlation of our instrument

[13]Without added controls Frankel and Romer (1999) report an F-statistic of 98.01 for their predicted trade shares.

[14]In Figure C-1 and Figure C-2 in the Online Appendix, we drop each donor and recipient one at a time in the bilateral sample, aggregate the data to the country level, and rerun the first stage regression, ignoring the uncertainty from the bilateral stage. The estimates vary only within a narrow band.

[15]The variation of aid induced by changes in divided donor governments is likely to be higher for recipients with many active donors. To investigate this, we measure donor fragmentation by a Herfindahl index and the combined share of the three largest donors. We then interact predicted aid to GDP with a dummy indicating whether the recipient has a higher donor fragmentation than the sample mean. The coefficients on predicted aid to GDP and the first stage F-statistics are qualitatively similar to those reported here. The interaction term is always insignificant.

with these variables might weaken the strength of our instrument and could violate the exclusion restriction in some circumstances. However, note that a violation of the exclusion restriction requires not only that fractionalization-induced aid disbursements vary in tandem with other variables and that these variables determine conflict, it also requires that these other variables have heterogeneous effects on regular and irregular aid recipients.[16]

C. Baseline results

We focus on a basic set of controls in our main specifications but allow for (fixed) unobserved country heterogeneity, unobserved time effects, and instrument our time-varying variable of interest. All of these three measures take care of omitted variables and contemporaneous endogeneity. We present two sets of estimates for our baseline results. Table III reports the regression results and Table IV shows the associated average partial effects of aid on the different transition probabilities.

Consider the regressions in Table III first. Column (1) shows the estimates without additional controls, column (2) controls for GDP and population, and the last column includes heterogeneous trends depending on the initial conflict state (by controlling for $t \times \mathbf{h}_{i0}$). The results are interesting in a couple of respects. The coefficients of aid to GDP and its interactions with the lagged states are very similar across all three specifications. The regressions suggest *i)* that the intensifying effect of aid on conflict is stronger if the country experienced a small conflict in the year before, and *ii)* that the effect is not statistically different from the base level (i.e., peace in the previous year) for the two higher conflict intensities. We also find reasonably strong evidence suggesting that aid is endogenous. The residuals from the first stage have the opposite signs and similar magnitudes as the coefficients on the base level. This suggests that we would find no evidence of an effect of aid on conflict without correcting for endogeneity (and this is indeed the case). In control function methods, testing the null that the coefficient on the

[16] Table C-1 in the Online Appendix reports results with controls for voting alignment in the United Nation's General Assembly, trade openness, and FDI inflows. The first stage estimates remain highly significant and within the range of those presented in Table II.

residuals is zero corresponds to a Hausman test of endogeneity which does not depend on the bilateral or first stage. The reported bootstrap standard errors will therefore be on the conservative side. Nevertheless, we can always reject the null of no endogeneity at the 5 or 10% significance level.

We prefer column (2) since it accounts for scale effects (conflicts with more battle-related deaths occur in larger countries) and measures the net effect of higher aid intensity at a given income level. One might argue that this introduces a bad control issue. However, none of the coefficients on the selected time varying controls are significant. Given the similarity in our point coefficients between the different models, we do not think that this is a major problem. Most existing studies use pooled methods which rely on between-country differences. Given that recipient level CREs and conflict histories are included in all of our specifications, log GDP and log population do not seem to contribute much additional information. Note that we defer the discussion of the lagged states to the next subsection where we analyze the persistence and duration of conflicts at various intensities.

We have strong reasons to trust the estimates presented in Table III. We allow for quasi-fixed effects, first-order multi-state dynamics, and correct for contemporaneous heterogeneity. In theory, additional controls may help justifying the identifying assumptions regarding the instrument but there is no *ex ante* reason to expect that our estimates are still biased. Including more variables also comes at a cost. Each additional variable consumes several degrees of freedom due to how the unobserved heterogeneity is modeled. We find some evidence in favor of different trends depending on the initial conflict state in column (3). However, these types of trends do not violate our identification assumption, as they do not differ across regular and irregular aid recipients. Allowing for them hardly changes the estimates of interest and only adds more complexity. We return to the issue of parallel trends and additional controls later when discussing extensions of the model.

To assess the magnitude of the implied effects we have to turn to partial effects as opposed to estimated coefficients. Table IV reports estimates of the APEs for a one percentage point change in aid on the various transition probabilities (see eq. 3). Note that—by definition—each row sums to zero. Although all estimates above the diagonal are positive and those below negative, we find no statistically significant evidence in favor of an effect of aid on conflict when countries are entirely at peace or engaged in an armed conflict with more than 25 BDs.

Aid has significant adverse effects in volatile environments which are not entirely peaceful but also not (yet) fully engaged in armed conflict. There, more aid makes peace less likely and a continuation of small conflict or a transition to armed conflict more likely. A one percentage point increase in the ratio of foreign aid to GDP leads to an approximate 1.4 percentage point increase in the probability of transitioning from small conflict to armed conflict.[17] The same increase in aid also significantly increases the likelihood of remaining in a small conflict (also by about 1.4 percentage points) and makes a transition to peace much less likely (by about -2.9 percentage points).[18]

The size of this effect is best understood in conjunction with a typical change in aid flows. The average aid to GDP ratio in our sample is about 5% and the within standard deviation is also close to 5% (when we exclude recipients who receive more than half their GDP in foreign aid, e.g., Liberia 2008, Palau 1994, 1995). Mali, for example, experienced a one standard deviation increase in its aid to GDP ratio in 1994 when the share of aid to GDP increased from about 8% to 13%. At the same time, there was an escalation from small conflict to armed conflict. Consistent with this observation, our model predicts an increase in the probability of transitioning from small conflict to armed conflict of about 7 percentage points. Aid increases of this magnitude are rare. Only in about 3% of the sample they exceed five percentage points but changes around one percentage point are more common (about 14% of the sample). In Uganda, for example, aid increased by about one percentage point on two occasions (1981 and 2002). In both cases, the country experienced an escalation of conflict.

[17]We might be concerned that the effect of aid on the transition from small conflict to armed conflict is driven by a small subset of observations. There are about 50 switches supporting this estimate and more than 300 observations behind each of the two lower switches.

[18]The size of the estimated effects are also in line with recent estimates by Besley and Persson (2011b), Crost et al. (2014), and Nunn and Qian (2014). However, de Ree and Nillesen (2009) find that an increase in aid flows by 10% decreases the probability of continuation of conflict by about eight percentage points.

D. Persistence, state dependence and duration

Table V shows the average transition probabilities as they are predicted by our preferred specification. The diagonal of this matrix shows the predicted persistence rates and the off-diagonal elements are the escalation and deescalation probabilities, respectively. Note that we use persistence and continuation as synonyms, so that persistence is simply the estimated probability of remaining in a particular state. The matrix provides nearly all the terms needed to estimate state dependence as in eq. 4 apart from the weights.

We find strong evidence of state dependence in each of the four states, even after controlling for observed and unobserved heterogeneity. State dependence in armed conflict and civil war is moderately high and very similar. For both types of conflict, the sheer fact that a country finds itself in conflict implies that the probability of remaining in conflict rises by at least 30 percentage points. Comparing these estimates with the persistence probabilities shown on the diagonal is particularly instructive. State dependence accounts for the bulk of persistence in armed conflict and civil war, but much less so in small conflict and peace.

Taking a truly dynamic approach allows us to bridge another distinction that is often drawn in the conflict literature: event models versus duration models. First-order Markov models can be compared to discrete time duration models with a constant hazard rate (e.g., Cappellari and Jenkins, 2004).[19] The expected duration of peace is about five years. Most conflicts are relatively short-lived on average. Small conflicts last about 1.4 years, armed conflict about 1.7 years, and civil wars about 1.5 years. We are predicting conflicts that last longer than three years only after about the 95th percentile (and longer than five years after the 99th percentile). This may seem short compared to other findings in the literature but it is worth bearing in mind that we distinguish between different types of conflict that are often lumped together. A conflict cycle that goes from small over armed conflict to outright civil war and back is perfectly compatible with the duration typically found in the literature (e.g., Collier and Hoeffler, 2004).

[19] To see the equivalence, recall that the hazard rate is the probability that the current state will end, or $\Pr(T_i = t | T_i \geq t)$. The probability of exiting a particular state is geometrically distributed with $\Pr[T_i = t] = p_{ii}^{t-1}(1 - p_{ii})$. The expected survival time in state i is $\mathbb{E}[T_i] = 1/(1 - p_{ii})$ and the quantile function is $Q(r) = F^{-1}(r) = \ln(1 - r)/\ln(p_{ii})$ where r is the percentile of interest.

E. Serial correlation in ε_{1it}

In deriving eq. 9, we assumed dynamic completeness which implies that ε_{1it} and, by extension, u_{1it} are serially uncorrelated. Recall that $u_{1it} = \omega u_{2it} + \varepsilon_{1it} = \omega(\nu_{2it} - b_{2i}) + \varepsilon_{1it}$. Any serial correlation in the control function errors, ν_{2it}, will therefore spill over into the main equation and violate this assumption.[20]

This has two implications. First, applications of the proposed estimation approach should test for serial correlation in the first stage residuals by regressing $\hat{\nu}_{2it}$ on $\hat{\nu}_{2i,t-1}$ (without a constant) and report the corresponding results. Second, if the null of no serial correlation is rejected, then eq. 9 must be modified appropriately. We follow Giles and Murtazashvili (2013) by swapping ν_{2it} for $\nu^*_{2it} = \nu_{2it} - \varrho\nu_{2i,t-1}$ in $t > 1$ and $\nu^*_{2i1} = (1 - \varrho^2)^{\frac{1}{2}}\nu_{2i1}$ in the first period. The Cochrane-Orcutt adjustment eliminates the remaining serial correlation in the first stage errors, while the Prais-Winsten adjustment preserves the first period. The APEs, state dependence and transition probabilities are then specified in terms of $\hat{\nu}^*_{2it}$.

Table VI report the corresponding results. We find strong evidence of serial correlation. Using the same data from column (2) of Table III, a regression of the first stage residuals on lagged values yields a coefficient of about 0.6 with a cluster-robust t-statistic of 9.16. The subsequent correction for serial correlation, however, only has a limited impact on our main findings. The interaction of the small conflict with aid ($a_{2it} \times h_{1,i,t-1}$) is still highly significant and of comparable size. There is one noticeable change. We no longer observe a strong correction of the level effect of aid by including the control function residuals.[21]

F. Identification assumptions and falsification

Our local average partial effect compares the effects of politically induced differences in bilateral aid between regular and irregular aid recipients. This raises the question whether

[20]To see this, assume $u_{2it} = \phi u_{2i,t-1} + e_{2it}$ with $\phi \neq 0$ and $\mathrm{Var}(e_{2it}) = \sigma^2_{e_2}$. It is straightforward to show that $\mathrm{Cov}(\varepsilon_{1it}, \varepsilon_{1i,t-1}) = \omega^2\phi\mathrm{Var}(e_{2it}) > 0$ in all but the trivial case of no endogeneity ($\omega = 0$).

[21]Donor government fractionalization does not change every year, so that our instrument contains a lot of serial correlation by construction. The wider standard error on the level effect of aid also spills over into the estimation of other terms in the APE matrix.

the parallel trends assumption inherent in difference-in-difference approaches is satisfied, or if spurious non-linear trends are at work. Our identification strategy would be invalid if the time-varying component of our instrument is spuriously correlated with the time trend in conflict and the strength of this time trend depends on the regularity of aid receipts.[22] Put differently, our finding that aid leads to an escalation of conflict rests on the coincident timing of politically-induced aid flows and the observed conflict histories. Randomizing aid flows along various dimensions allows us to break this temporal structure.

Figure II reports the results from 999 Monte Carlo simulations for four randomization strategies. We randomly reassign the aid to GDP ratio by exchanging *i)* all observations in the sample—labelled *Overall*, *ii)* the entire time series between countries—labelled *Countries*, *iii)* years within countries—labelled *Within*, and *iv)* countries within years—labelled *Years*. For each placebo test, we report the distribution of the coefficients on the interaction terms. The p-values are calculated as the proportion of times that the absolute value of the t-statistics in the simulated data exceed the absolute value of the original t-statistic. The results unambiguously show that our findings are not driven by global trends, cross-sectional dependence, or selection of countries into regular aid receipts. While the original data generate a strong interaction effect of aid with the lag of small conflict, the simulations are centered on zero and very rarely include such an effect by chance.

G. Further tests of robustness

We run a battery of robustness checks to verify our main findings (all corresponding results are relegated to Table D-1 to Table D-9 in the Online Appendix and discussed there in detail). The results are robust to *i)* specifying linear models to show that we are not identifying these effects by functional form only, *ii)* changing the outcome variable to variants of the industry standard, which illustrates that the modelling of small conflicts is key, *iii)* leaving out particular types of violence in the definition of 'small conflict', *iv)* altering the definition of foreign aid, *v)* adding controls in the instrument

[22] See Christian and Barrett (2017) who show that such non-linear trends could be driving the positive effect of food aid on conflict documented in Nunn and Qian (2014).

generation stage to account for bilateral correlations in trade and political alignments, *vi)* including additional covariates, such as neighbors in conflict, in the main model, and, *vii)* considering the roles played by multilateral and humanitarian aid. Last but not least, we compare our findings to those presented in Nunn and Qian (2014) and present falsification tests using their instruments.

6. Conclusion

This paper studies the effects of development aid on conflict. While there is a large literature on the topic, it typically separates the onset of a conflict from its continuation and neglects smaller acts of violence. This misses important dynamics which our paper makes an effort to expose. We introduce a dynamic ordered probit framework which can account for unobserved heterogeneity and endogenous variables, together with an identification strategy based on the characteristics of the electoral system of donor countries.

Our results show that the effects of bilateral aid are heterogeneous with respect to the different intensity levels of conflict. Whereas aid increases the probability that a conflict escalates from a low level of political violence to armed conflict, we find little evidence in favor of an adverse effect of aid in truly peaceful countries. Aid does also not seem to affect the transition probabilities once a country experiences armed conflict or civil war. These results underline the importance of separating truly peaceful situations from countries exposed to small conflict. If we would not account for this distinction, we would fail to detect an effect of aid on conflict.

These findings call for care when devising aid policies for countries affected by conflict. Particular care has to be exercised when aid is given to countries where turmoil is already present but armed conflict has not yet erupted. Our results suggest that aid might be more harmful than helpful in these situations, despite best intentions. Our analysis focuses on conflict but the empirical framework we offer could be useful in many other settings with ordered outcomes.

References

Acemoglu, D. and A. Wolitzky (2014). Cycles of conflict: An economic model. *American Economic Review 104*(4), 1350–1367.

Ahmed, F. Z. and E. Werker (2015). Aid and the rise and fall of conflict in the Muslim world. *Quarterly Journal of Political Science 10*(2), 155–186.

Banks, A. S. and K. A. Wilson (2015). *Cross-National Time-Series Data Archive.* Jerusalem: Databanks International.

Bazzi, S. and C. Blattman (2014). Economic shocks and conflict: Evidence from commodity prices. *American Economic Journal: Macroeconomics 6*(4), 1–38.

Beck, N., J. N. Katz, and R. Tucker (1998). Taking time seriously: Time-series-cross-section analysis with a binary dependent variable. *American Journal of Political Science 42*(4), 1260–1288.

Beck, T., G. Clarke, A. Groff, P. Keefer, and P. Walsh (2001). New tools in comparative political economy: The database of political institutions. *World Bank Economic Review 15*(1), 165–176.

Becker, G. S. (1968). Crime and punishment: An economic approach. *Journal of Political Economy 76*(2), 169–217.

Berman, N. and M. Couttenier (2015). External shocks, internal shots: The geography of civil conflicts. *Review of Economics and Statistics 97*(4), 758–776.

Besley, T. and T. Persson (2011a). Fragile states and development policy. *Journal of the European Economic Association 9*(3), 371–398.

Besley, T. and T. Persson (2011b). The logic of political violence. *Quarterly Journal of Economics 126*(3), 1411–1445.

Blattman, C. and E. Miguel (2010). Civil war. *Journal of Economic Literature 48*(1), 3–57.

Brückner, M. (2013). On the simultaneity problem in the aid and growth debate. *Journal of Applied Econometrics 28*(1), 126–150.

Bueno de Mesquita, E. (2013). Rebel tactics. *Journal of Political Economy 121*(2), 323–357.

Cappellari, L. and S. P. Jenkins (2004). Modelling low income transitions. *Journal of Applied Econometrics 19*(5), 593–610.

Christian, P. and C. B. Barrett (2017). Revisiting the effect of food aid on conflict: A methodological caution. Policy Research Working Paper 8171, World Bank.

Collier, P. and A. Hoeffler (2004). Greed and grievance in civil war. *Oxford Economic Papers 56*(4), 563–595.

Crost, B., J. Felter, and P. Johnston (2014). Aid under fire: Development projects and civil conflict. *American Economic Review 104*(6), 1833–1856.

de Ree, J. and E. Nillesen (2009). Aiding violence or peace? The impact of foreign aid on the risk of civil conflict in Sub-Saharan Africa. *Journal of Development Economics 88*(2), 301–313.

Dreher, A. and A. Fuchs (2011). Does terror increase aid? *Public Choice 149*(3-4), 337–363.

Dreher, A., A. Fuchs, R. Hodler, B. Parks, P. A. Raschky, and M. J. Tierney (2019). African leaders and the geography of China's foreign assistance. *Journal of Development Economics 140*, 44–71.

Dreher, A. and S. Langlotz (2020). Aid and growth. New evidence using an excludable instrument. *Canadian Journal of Economics*, forthcoming.

Dube, O. and S. Naidu (2015). Bases, bullets and ballots: The effect of U.S. military aid on political conflict in Colombia. *Journal of Politics 77*(1), 249–267.

Esteban, J., L. Mayoral, and D. Ray (2012). Ethnicity and conflict: An empirical study. *American Economic Review 102*(4), 1310–1342.

Fearon, J. D. and D. D. Laitin (2003). Ethnicity, insurgency, and civil war. *American Political Science Review 97*(1), 75–90.

Frankel, J. A. and D. Romer (1999). Does trade cause growth? *American Economic Review 89*(3), 379–399.

Giles, J. and I. Murtazashvili (2013). A control function approach to estimating dynamic probit models with endogenous regressors. *Journal of Econometric Methods 2*(1), 69–

87.

Gleditsch, N. P., P. Wallensteen, M. Eriksson, M. Sollenberg, and H. Strand (2002). Armed conflict 1946-2001: A new dataset. *Journal of Peace Research 39*(5), 615–637.

Grossman, H. I. (1991). A general equilibrium model of insurrections. *American Economic Review 81*(4), 912–921.

Nunn, N. and N. Qian (2014). U.S. food aid and civil conflict. *American Economic Review 104*(6), 1630–1666.

Persson, T., G. Roland, and G. Tabellini (2007). Electoral rules and government spending in parliamentary democracies. *Quarterly Journal of Political Science 2*(2), 155–188.

Rabe-Hesketh, S. and A. Skrondal (2013). Avoiding biased versions of Wooldridge's simple solution to the initial conditions problem. *Economics Letters 120*(2), 346–349.

Rivers, D. and Q. H. Vuong (1988). Limited information estimators and exogeneity tests for simultaneous probit models. *Journal of Econometrics 39*(3), 347–366.

Rohner, D., M. Thoenig, and F. Zilibotti (2013). War signals: A theory of trade, trust, and conflict. *Review of Economic Studies 80*(3), 1114–1147.

Wooldridge, J. M. (2005). Simple solutions to the initial conditions problem in dynamic, nonlinear panel data models with unobserved heterogeneity. *Journal of Applied Econometrics 20*(1), 39–54.

Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data* (2nd ed.). Cambridge (MA): MIT Press.

Tables and figures

Table I
Unconditional transition matrix (in %)

	To State			
From State	Peace	Small Conflict	Armed Conflict	Civil War
Peace	87.26	10.69	2.06	0.00
Small Conflict	43.85	48.13	6.78	1.24
Armed Conflict	11.28	8.46	70.30	9.96
Civil War	1.49	5.97	23.88	68.66

Notes: The table reports the raw transition matrix estimated using the same balanced sample of 125 countries over 36 years that is used in the main analysis (4,500 observations imply 4,375 transitions). Rows sum to 100%.

Table II
First stage regressions with generated IV

	Dependent Variable: Aid to GDP		
Variables	(1)	(2)	(3)
Predicted aid to GDP ($\sum_j \hat{a}_{3ijt}$)	1.3520***	1.2326***	1.2232***
	(0.1037)	(0.0907)	(0.0863)
Selected Controls			
Log GDP		-5.1140***	-5.0876***
		(0.8395)	(0.8453)
Log Population		6.0835***	5.8781**
		(2.2943)	(2.3769)
Additional Controls			
Country FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Initial states × linear trend	No	No	Yes
Summary Statistics			
Kleibergen-Paap F-statistic IV	170.1	184.9	200.8
$N \times T$	4,375	4,375	4,375
T	35	35	35
N	125	125	125
Within-R^2	0.0412	0.0763	0.0782

Notes: The table shows the results of first stage regressions using a linear two-way fixed effects model. The instrument is the sum of predicted bilateral aid over all donors ($\sum_j \hat{a}_{3ijt}$) from eq. 11. Two-step panel bootstrap standard errors—sampling over recipients in the bilateral sample and computed with 999 replications—are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table III
Three-step ordered probit regressions, CRE and CF

	Dependent Variable: Ordered Conflict		
Variables	(1)	(2)	(3)
Aid to GDP (a_{2it})	0.0728*	0.0721	0.0717
	(0.0410)	(0.0443)	(0.0492)
Residuals ($\hat{\nu}_{2it}$)	-0.0847**	-0.0863**	-0.0855*
	(0.0407)	(0.0439)	(0.0485)
Interactions with Lagged States			
Small Conflict ($a_{2it} \times h_{1,i,t-1}$)	0.0220***	0.0212**	0.0209**
	(0.0083)	(0.0090)	(0.0092)
Armed Conflict ($a_{2it} \times h_{2,i,t-1}$)	-0.0084	-0.0106	-0.0104
	(0.0190)	(0.0200)	(0.0192)
Civil War ($a_{2it} \times h_{3,i,t-1}$)	-0.0023	-0.0023	-0.0020
	(0.0233)	(0.0241)	(0.0239)
Lagged States			
Small Conflict ($h_{1,i,t-1}$)	0.5823***	0.5760***	0.5597***
	(0.0777)	(0.0792)	(0.0779)
Armed Conflict ($h_{2,i,t-1}$)	2.1102***	2.1068***	2.0683***
	(0.1816)	(0.1833)	(0.1980)
Civil War ($h_{3,i,t-1}$)	3.4290***	3.4245***	3.3762***
	(0.2275)	(0.2192)	(0.2360)
Additional Controls			
Controls	No	Yes	Yes
Initial values × trend	No	No	Yes
Recipient CRE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Residual CRE	Yes	Yes	Yes
Initial states	Yes	Yes	Yes
Summary Statistics			
$N \times T$	4,375	4,375	4,375
T	35	35	35
N	125	125	125

Notes: The table shows the results of an ordered probit model with correlated random effects and a control function approach. Three-step panel bootstrap standard errors—sampling over recipients in the bilateral sample and computed with 999 replications—are in parentheses. All models also estimate J cut points and the variance of the random recipient effect (not reported). Controls refer to the log of GDP and the log of population. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table IV

Average partial effect of aid on transition probabilities

	To State			
From State	Peace	Small Conflict	Armed Conflict	Civil War
Peace	-1.6333	1.1389	0.4835	0.0108
	(1.1138)	(0.7759)	(0.3402)	(0.0123)
Small Conflict	-2.8186**	1.4175**	1.3305**	0.0706
	(1.4156)	(0.7165)	(0.6776)	(0.0576)
Armed Conflict	-1.3965	-0.4973	1.3356	0.5583
	(1.2676)	(0.4808)	(1.1699)	(0.5224)
Civil War	-0.4252	-0.9872	-0.5211	1.9334
	(0.4325)	(0.7791)	(0.5410)	(1.5116)

Notes: The table reports the average partial effect of aid on the different transition probabilities. The estimates are based on column (2) in Table III. Three-step panel bootstrap standard errors—sampling over recipients in the bilateral sample and computed with 999 replications—are in parentheses. Rows sum to zero. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table V

Estimated transition probabilities and state dependence

	To State			
From State	Peace	Small Conflict	Armed Conflict	Civil War
Peace	79.9540***	16.3445***	3.6569***	0.0447*
	(1.9269)	(1.4555)	(0.7171)	(0.0258)
Small Conflict	61.7511***	27.4629***	10.4962***	0.2899**
	(2.4831)	(1.8718)	(1.3449)	(0.1337)
Armed Conflict	21.7830***	32.6899***	39.7492***	5.7778***
	(3.8680)	(2.0595)	(4.3945)	(1.2866)
Civil War	3.4853*	13.8347***	51.1025***	31.5775***
	(1.9110)	(3.0426)	(3.5621)	(4.3390)
State Dependence	40.7941***	8.8896***	32.3801***	30.7649***
	(2.5223)	(1.4512)	(4.2788)	(4.2655)

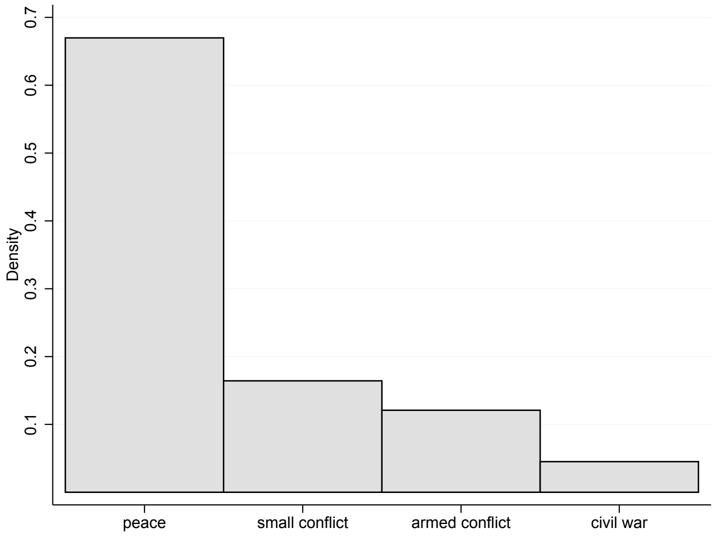
Notes: The table reports the estimated transition probabilities and state dependence. The estimates are based on column (2) in Table III. Three-step panel bootstrap standard errors—sampling over recipients in the bilateral sample and computed with 999 replications—are in parentheses. The upper four rows sum to 100%. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table VI
Correcting for serial correlation

Variables	*Dependent Variable: Ordered Conflict* (1)	(2)	(3)
Aid to GDP (a_{2it})	-0.0145	-0.0144	-0.0142
	(0.0099)	(0.0102)	(0.0106)
Residuals ($\hat{\nu}_{2it}$)	0.0078	0.0040	0.0039
	(0.0087)	(0.0092)	(0.0086)
Interactions with Lagged States			
Small Conflict ($a_{2it} \times h_{1,i,t-1}$)	0.0201**	0.0194**	0.0190**
	(0.0079)	(0.0089)	(0.0090)
Armed Conflict ($a_{2it} \times h_{2,i,t-1}$)	-0.0089	-0.0119	-0.0118
	(0.0195)	(0.0203)	(0.0197)
Civil War ($a_{2it} \times h_{3,i,t-1}$)	-0.0039	-0.0038	-0.0035
	(0.0231)	(0.0243)	(0.0241)
Lagged States			
Small Conflict ($h_{1,i,t-1}$)	0.5906***	0.5830***	0.5659***
	(0.0775)	(0.0799)	(0.0786)
Armed Conflict ($h_{2,i,t-1}$)	2.1190***	2.1183***	2.0792***
	(0.1824)	(0.1820)	(0.1973)
Civil War ($h_{3,i,t-1}$)	3.4400***	3.4350***	3.3867***
	(0.2259)	(0.2165)	(0.2332)
Additional Controls			
Controls	No	Yes	Yes
Initial value × trend	No	No	Yes
Recipient CRE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Residual CRE	Yes	Yes	Yes
Initial States	Yes	Yes	Yes
Summary Statistics			
$N \times T$	4,375	4,375	4,375
T	35	35	35
N	125	125	125

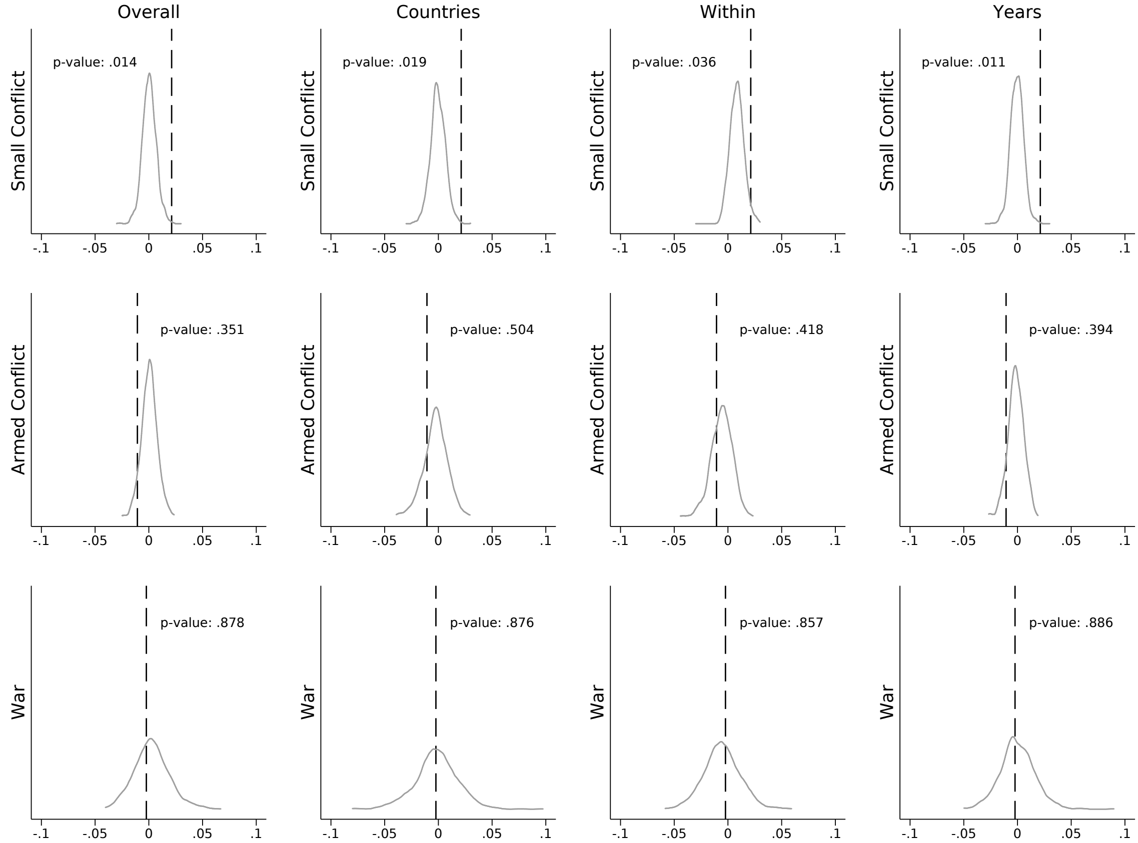
Notes: The table shows the results of an ordered probit model with correlated random effects and a control function approach with a correction for serial correlation in the reduced form errors. Three-step panel bootstrap standard errors—sampling over recipients in the bilateral sample and computed with 999 replications—are in parentheses. All models also estimate J cut points and the variance of the random recipient effect (not reported). Controls refer to the log of GDP and the log of population. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Figure I
Distribution of conflict intensities



Notes: Illustration of the unconditional distribution of the ordinal conflict measure. There are 3,014 peace years, 739 small conflict years, 544 armed conflict years, and 203 civil war years in our sample.

Figure II
Randomization Test



Notes: The figure shows the distribution of point coefficients of our main interaction terms based on 999 Monte Carlo replications. '*Overall*' randomizes all observations, '*Countries*' entire time series between countries, '*Within*' years within countries, and '*Years*' countries within years. Dashed vertical lines indicate the original estimate from column (2) in Table III. The p-values are calculated as the proportion of times that the absolute value of the t-statistics in the simulated data exceed the absolute value of the original t-statistic.

Supplementary Online Appendix to 'Fueling Conflict? (De)Escalation and Bilateral Aid' by Richard Bluhm, Martin Gassebner, Sarah Langlotz and Paul Schaudt

A. Countries and summary statistics

This section provides additional details and summary statistic of the sample underlying the analysis in the main text.

List A-1 provides a list of the 28 included donor countries.

List A-2 provides a list of the 125 recipients. The balancing requirement implies that we omit some countries for which we do not have consistent data over the entire period from 1975 to 2010.

Table A-1 adds summary statistics of the main variables used in the analysis. Panel A shows statistics for the full (unbalanced) bilateral data. Panel B shows statistics for the (balanced) country-level data.

List A-1

Included donor countries, in alphabetical order

Australia, Austria, Belgium, Canada, Czech Republic, Denmark, Finland, France, Germany, Greece, Iceland, Ireland, Italy, Japan, Korea, Luxembourg, Netherlands, New Zealand, Norway, Poland, Portugal, Slovak Republic, Slovenia, Spain, Sweden, Switzerland, United Kingdom, United States.

List A-2

Included recipient countries, in alphabetical order

Afghanistan, Albania, Algeria, Angola, Antigua and Barbuda, Argentina, Bahamas, Bahrain, Bangladesh, Barbados, Belize, Benin, Bhutan, Bolivia, Botswana, Brazil, Brunei Darussalam, Burkina Faso, Burundi, Cambodia, Cameroon, Cape Verde, Central African Republic, Chad, Chile, China, Colombia, Comoros, Congo, Costa Rica, Cuba, Cyprus, Democratic Republic of Congo, Djibouti, Dominica, Dominican Republic, Ecuador, Egypt, El Salvador, Equatorial Guinea, Ethiopia, Fiji, Gabon, Gambia, Ghana, Grenada, Guatemala, Guinea, Guinea-Bissau, Guyana, Haiti, Honduras, India, Indonesia, Iran, Iraq, Israel, Ivory Coast, Jamaica, Jordan, Kenya, Kiribati, Lao, Lebanon, Lesotho, Liberia, Madagascar, Malawi, Malaysia, Maldives, Mali, Malta, Marshall Islands, Mauritania, Mauritius, Mexico, Micronesia, Mongolia, Morocco, Mozambique, Namibia, Nepal, Nicaragua, Niger, Nigeria, Oman, Pakistan, Palau, Panama, Papua New Guinea, Paraguay, Peru, Philippines, Rwanda, Saint Kitts and Nevis, Saint Lucia, Saint Vincent and the Grenadine, Samoa, Sao Tome and Principe, Senegal, Seychelles, Sierra Leone, Singapore, Solomon Islands, Somalia, South Africa, Sri Lanka, Sudan, Suriname, Swaziland, Syria, Tanzania, Thailand, Togo, Tonga, Trinidad and Tobago, Tunisia, Turkey, Uganda, Uruguay, Vanuatu, Venezuela, Vietnam, Zambia, Zimbabwe.

Table A-1
Summary statistics

Variables	Mean	Std. Dev.	Min	Max	N
Panel A) Bilateral					
Aid to GDP (in percent)	0.19	1.40	-5.68	228.67	131,964
Government Fractionalization	0.30	0.27	0.00	0.83	141,789
Legislative Fractionalization (FPTP only)	0.06	0.17	0.00	0.69	151,906
Probability to receive Aid	0.46	0.37	0.00	1.00	152,208
Panel B) Country-level					
Aid to GDP (in percent)	4.95	8.84	-2.95	241.69	4,500
Log of GDP	16.19	2.10	11.39	22.97	4,500
Log of Population	8.17	2.24	2.50	14.11	4,500

Notes: The aid to GDP ratio has a maximum well in excess of 200%. This maximum is driven by Palau. Together with other pacific islands, Palau is part of the Compact of Free Association with the United States and receives foreign assistance greatly exceeding its GDP. Without Palau, the maximum falls to slightly above 100% (due to Liberia). Negative numbers are repayments of loans.

B. The ordinal measure in Sri Lanka

This section illustrates the dynamics of the civil conflict in Sri Lanka from 1975 to 2010 as captured by our measure. Sri-Lanka is an ideal case for two reasons: First, the conflict went through all conflict intensities. Second, the conflict turned violent in the mid-1970s right around the start of our sample and ended in 2010 at the end of our sample period.

Figure B-1 tracks this evolution using our measure and highlights significant events. The political conflict between the Sinhalese (about 73.8% of the population) and the Tamils (about 18% of the population, concentrated in the northeast of the country), has been lingering in Sri Lanka since the independence from the British Empire in 1948. The conflict started escalating in 1970 when the new constitution declared Sinhala as the official language and defined Buddhism as the official religion. The reaction of the Tamil (mainly Christians and Hindus with their own language) followed in 1972 when Ceylon became officially recognized as the Republic of Sri Lanka.[1] The Tamils formed the Tamil New Tigers Group to set up a separate homeland *Tamil Eelam* in the northeast of Sri Lanka which was accompanied by heavy riots (Banks and Wilson, 2015).[2]

In 1975, the New Tigers Group re-named itself the Liberation Tigers of Tamil Eelam (LTTE) spurring harsh responses by the government. Notice that while the UCDP-PRIO still codes the country as peaceful, our residual category of small conflicts already picks up the escalating violence. In 1978 the LTTE was outlawed. Interestingly, this coincides with a drop in our conflict measure to zero. The next escalation occurred in 1981, when riots erupted in Jaffna and a state of emergency was declared. Finally, in 1983 the first guerrilla attack, an ambush, was conducted by the LTTE, resulting in the death of 13 soldiers. The incident led to the eruption of riots and the killing of hundreds of people. The year 1984 then marks the first armed conflict observation in the UCDP-PRIO data set (category two in our measure).

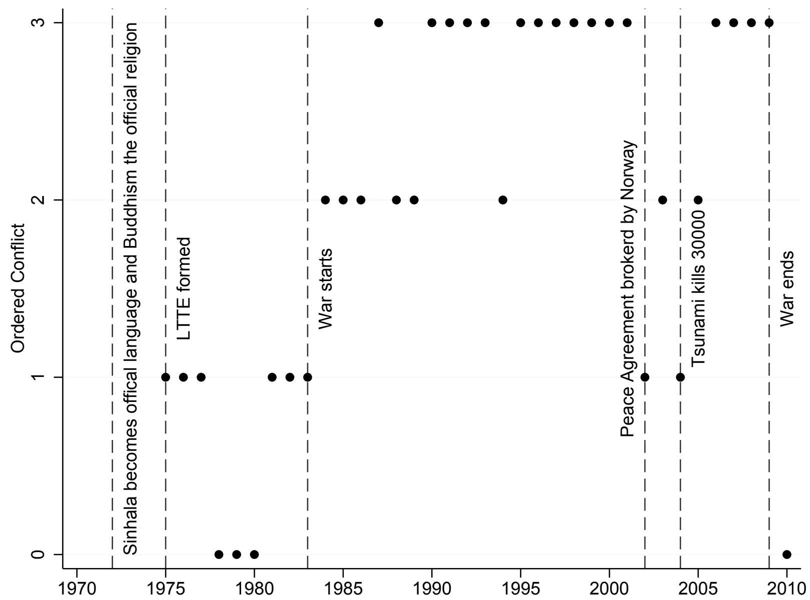
The UCDP-PRIO data set does a good job for most of the following years in which the conflict is varying between armed conflict and civil war until the military defeat of the LTTE in 2009.[3] There are, however, two observations, one in 2002 and the other in 2004, in which UCDP-PRIO codes a peace observation. In both cases what follows is an armed conflict observation, and in 2006 a civil war observation. The two "peace" observations which in our approach fall into the small conflict category coincide the ceasefire mediated by Norway in 2002 and the split of LTTE, after which one part formed a pro-government party. The second slump in conflict intensity was 2004, in which more than 30,000 citizens died during the tsunami.[4] Yet in both cases violence never ceased but failed to reach the threshold of 25 BD. In 2002 there have still been several clashes between LTTE fighters

[1]See http://www.cfr.org/terrorist-organizations-and-networks/sri-lankan-conflict/p11407.
[2]See http://www.aljazeera.com/focus/blanktemplate/2008/11/2008111061193133.html.
[3]See http://www.nytimes.com/2009/05/19/world/asia/19lanka.html?_r=2&ref=global-home.
[4]See http://www.cfr.org/terrorist-organizations-and-networks/sri-lankan-conflict/p11407.

and government soldiers, although both groups tried to adhere to the peace agreement.[5] In 2004 rioters burned down outlets of the government friendly splinter group who seceded from the LTTE (Banks and Wilson, 2015).[6]

Summing up, our measure captures the cyclical nature of the civil conflict between the LTTE and the government of Sri-Lanka rather well. Sri-Lanka was never actually completely at peace from 1981–2009 until the military defeat of the LTTE.

Figure B-1
Conflict dynamics in Sri Lanka



[5]See http://www.hiik.de/en/konfliktbarometer/pdf/ConflictBarometer_2002.pdf.
[6]See http://www.hiik.de/en/konfliktbarometer/pdf/ConflictBarometer_2004.pdf.

C. Robustness of the first stage relationship

This section probes the strength and robustness of the first stage relationship.

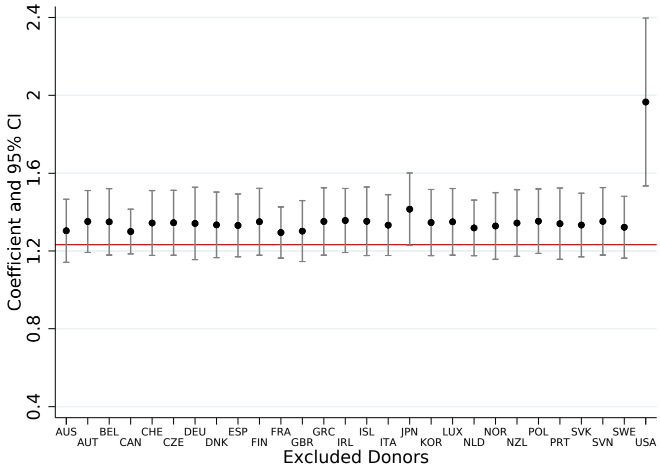
Figure C-1 investigates the influence each of the 28 donors individually has on the strength of the first stage regression by dropping them one at a time, aggregating aid without this donor, and re-estimating the first stage relation. Note that we ignore the uncertainty introduced by estimating the bilateral stage for this exercise. Removing any of the donors, apart from the U.S., only has a limited influence on the estimates. Removing the U.S. leads to a considerably larger first stage coefficient. Political fractionalization seems to underpredict politically-induced aid flows more strongly in the remaining 27 donors, so that these predictions have to be scaled up more. Nevertheless, the first stage without the U.S. is still very strong (with a cluster-robust F-statistic of 79.83).

Figure C-2 conducts a similar exercise but drops recipients instead of donors. Here we observe very little variation. The estimates remain close to the full sample in each instance.

Table C-1 expands the first stage regressions by including voting alignment in the UN General Assembly (based on the ideal points estimated in Bailey et al., 2017), trade openness, and FDI inflows over GDP as additional controls. The overall strength of our instrument is virtually unaffected but varies strongly with the size of the panel. The F-statistic of the instrument varies between 50 to 200. Likewise, the estimated coefficients on predicted aid remain in the range of 1.2 to 1.35. Closer voting alignment and more openness increase aid flows, while the coefficient on FDI flows is not significant at conventional levels. While these measures clearly matter for aid allocation, they do not capture the exogenous variation that is contained in our instrument.

Figure C-1

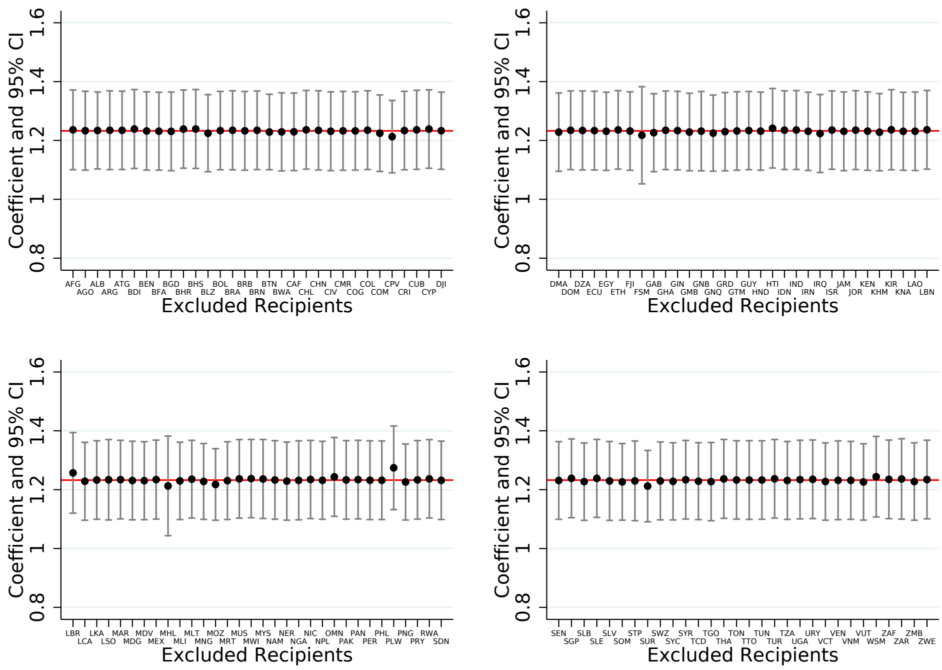
Leave-one-out test: Donors



Notes: Each point in the figure represents the result of a regression of actual on predicted aid shares where one of the DAC donors has been excluded from the bilateral sample. 95% confidence intervals based on standard errors clustered on the recipient are indicated as error bars. The red line indicates the original estimate from column (2) in Table II using all donors.

Figure C-2

Leave-one-out test: Recipients



Notes: Each point in the figure represents the result of a regression of actual on predicted aid shares where one of the recipients has been excluded from the bilateral sample. 95% confidence intervals based on standard errors clustered on the recipient are indicated as error bars. The red line indicates the original estimate from column (2) in Table II using all recipients.

Table C-1
First stage with additional controls

Variables	Dependent Variable: Aid to GDP (1)	(2)	(3)	(4)
Predicted aid to GDP ($\sum_j \hat{a}_{3ijt}$)	1.3473***	1.2166***	1.3761***	1.3076***
	(0.1916)	(0.0899)	(0.1386)	(0.1930)
Selected Controls				
Log GDP	-3.9135***	-5.5708***	-4.8307***	-4.1218***
	(0.9511)	(0.9404)	(0.9123)	(0.9176)
Log Population	6.0154**	5.8195***	7.0225***	6.4802***
	(2.5290)	(2.1925)	(2.1647)	(2.4672)
UNGA Alignment	1.9791***			1.7050***
	(0.5616)			(0.5080)
Trade Openness		0.0309**		0.0408***
		(0.0155)		(0.0092)
FDI Inflows to GDP			0.0186	0.0198
			(0.0259)	(0.0293)
Additional Controls				
Country FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
Summary Statistics				
Kleibergen-Paap F-statistic IV	49.47	183.1	98.64	45.89
Within-R^2	0.142	0.0842	0.118	0.175
$N \times T$	3,150	4,375	4,200	3,080
T	35	35	35	35
N	90	125	120	88

Notes: The table shows the results of first stage regressions using a linear two-way fixed effects model. The instrument is the sum of predicted bilateral aid over all donors ($\sum_j \hat{a}_{3ijt}$) from eq. 11. Two-step panel bootstrap standard errors—sampling over recipients in the bilateral sample and computed with 999 replications—are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

D. Additional regression results

We present a number of extensions which subject our main findings to several robustness checks and perturbations. First, we compare the ordered probit estimator to standard linear models. Second, we examine the sensitivity of our results to the underlying definition of the key variables. Third, we include a variety of additional controls. Finally, we examine the role of multilateral and humanitarian aid.

Linear estimation: The proposed dynamic ordered probit model is reasonably demanding to estimate and one might be concerned that our findings are driven by the structure we impose on the data. Table D-1 addresses this issue. Here we ignore the ordinal nature and estimate our base specification using different linear approaches. Recall that least squares is not suitable for ordinal outcomes if the number of outcomes is not large and the error distribution is not approximately normal, among other issues. We ignore the uncertainty introduced in the bilateral stage when estimating the linear models, as the parameters are estimated consistently and the limiting distributions are the same (see Wooldridge, 2010, p. 125)

All first order effects of aid on conflict are similar to the non-linear models. Column (1) in Table D-1 shows that, just as in the non-linear models, we find no effect if we estimate the fixed effects OLS counterpart to our dynamic specification when ignoring the endogeneity of aid. Column (2) then uses a control function approach to correct for the endogeneity of aid and recovers a positive first order effect of aid on all conflict outcomes. Column (3) illustrates the well-known equivalence of control function (CF) and instrumental variables (IV) approaches (which is approximate here since we do not control for the initial states in the CF approach, mirroring the ordinal model proposed in the main text).

The models with interaction terms confirm our main findings. As columns (5) and (6) show, once we correct for the endogeneity of aid, the estimated coefficient is positive and significant. The coefficients on the three interaction terms are numerically similar, no matter if we use the control function estimator or not. In column (6), when we use a standard IV approach, the interaction effects become much less precisely estimated while the signs and magnitudes are broadly stable. The CF estimator requires only one first stage estimation to correct for popular transformations (such as squares or interactions) of the endogenous variable. The IV estimator instead requires us to generate many additional instruments to run as many additional first stage regression as we have interaction terms. As a result, the IV estimator is much less efficient but imposes fewer assumptions (Wooldridge, 2010, pp. 128–129). Given the stability of the estimated coefficients, this difference appears to be immaterial in our case.[7]

[7] A drawback of the IV results is that the coefficients on the lagged states suffer from Nickell bias

Definition of variables: We now summarize the sensitivity of our results with respect to the operationalization of our key variables. In Table D-2 we alter the construction of our conflict and aid measures. Column (1) addresses the potential concern that while our newly developed measure is a step forward, we might not have gone far enough. One type of violence which we have so far neglected is terrorism. In times of major civil conflicts the definition of what constitutes a terrorist act becomes very blurred. In fact Campos and Gassebner (2013) show that countries with a history of civil wars are the ideal training ground for (international) terrorists. We now include country-year observations with a positive number of terror attacks[8] but less than 25 BDs in the category one (*small conflict*) of our ordinal measure. In column (2) we combine categories two and three, since several studies only distinguish between peaceful countries and countries with more than 25 BDs. In both cases the results are qualitatively similar to our main findings.

Next, we compare our approach to the industry standard, where peace and small conflict are combined in one category. This eliminates the possibility to distinguish between truly peaceful countries and countries that experience small conflict. In line with our expectations, neither the level estimates nor the interaction effects are statistically significant in column (3). This is also true for the APEs.

Our results are not driven by one single dimension of small conflict. We code variants of the small conflict category by excluding one of the constituting variables each time (e.g., riots, assassinations). As Table D-3 shows, we obtain very similar results for all perturbations apart from riots.

Table D-4 changes the definition of aid. So far, we have only focused on net ODA. In column (1) we test the effect of gross ODA. We strictly prefer net ODA since it takes repayments into account, but the results remain qualitatively similar given the differences in scale. In column (2) we include OOF to capture a broader concept of financial inflows from abroad, again the results remain the same. In columns (3) and (4) we exclude Canada, the UK and the U.S. (the biggest and potentially most politically-motivated donor in the world, e.g., Kuziemko and Werker, 2006, among many others). We do so for two reasons. First, for those three countries we use legislative fractionalization rather than government fractionalization as an IV for bilateral aid. In order to rule out that our results depend on this choice, we estimate our preferred specification for the remaining 25 DAC donors. Second, these three donors could differ from the rest of the DAC donors in how they disburse aid to countries in conflict (e.g., if they are important to the U.S.). Column (3) uses ODA, while Column (4) uses ODA with OOF. In each case, the estimated coefficients are in line with our preferred specification.

which, due to the presence of six interdependent lagged states and interactions, may not be trivial even with moderately large T.

[8]From START (National Consortium for the Study of Terrorism and Responses to Terrorism) (2016).

Bilateral correlation: Another worry might be that bilateral correlation in trade or political proximity is correlated with aid flows and, hence, inadvertently part of the constructed instrument. Table D-5 adds trade openness, voting alignment in the UN General Assembly (based on the ideal points estimated in Bailey et al., 2017), or both to the bilateral relationship. While bilateral trade comes out very strongly, the estimated coefficient on the interaction of government fractionalization with the probability to receive aid is very similar to before. The coefficient on the interaction of legislative fractionalization changes substantially, but it is only informed by three donor countries and the sample size is substantially smaller. We are ultimately concerned with whether this changes our main results. Table D-6 reproduces our baseline regression but partials out the effects of these bilateral variables. The interaction of aid to GDP with small conflict remains significant but the control function residuals no longer indicate a strong correction for endogeneity.

Additional controls: In Table D-7 we extend the set of control variables. Column (1) examines influence of conflict in the immediate regional neighborhood. We find little evidence of spillover effects, although such peer effects are generally difficult to identify. This is in line with Bosker and de Ree (2014), who find that only ethnic conflicts spill across borders. Columns (2) to (5) examine if political institutions affect the link between aid and conflict. This comes at the cost of a reduced sample.[9] None of the political variables alter our main results. Political instability is associated with conflict and countries with a Polity IV score of greater or equal to six are less likely to engage in violent activities. Column (6) shows that GDP growth makes conflict less likely but does not affect the relationship between aid and conflict.

We strongly prefer our baseline estimates with country and time effects over the results reported in Table D-7. Many of the added variables can be considered "bad controls" in the sense that they themselves could be outcomes of development aid. As cases in point, political instability, classification as a democracy, or GDP growth have all been causally linked to aid in the past. The inclusion of outcomes on the right hand side creates a selection problem which can completely distort the estimated causal effect.[10]

Multilateral, humanitarian, and food aid: Multilateral aid is typically a bit less than one third of all aid. To estimate its influence, we first calculate the correlation of multilateral aid as a share of GDP with aggregated predicted aid to GDP (our instrument)

[9] The Polity IV score is not available for cases of foreign "interruption" (code -66) and lacks data for island countries. We lose, e.g., Afghanistan, Iraq, Cambodia, and Lebanon.

[10] See Angrist and Pischke (2008) for a discussion of this problem. A similar reasoning could be used to prefer the short specification in column (1) of Table III over the other two columns. Note that the inclusion of log GDP and log population hardly makes a difference in the estimates and both variables have insignificant coefficients, so that this distinction is immaterial for our main results.

and then the correlation with the part of our instrument that is solely driven by exogenous variation.[11] The correlation of multilateral aid to GDP with aggregated predicted aid to GDP is 0.46, but falls to 0.05 when the exogenous component is isolated. Hence we conclude that multilateral aid is certainly important and correlates with bilateral aid but not with our identifying variation.

Next, we consider the role of humanitarian and—its main component—food aid. Although humanitarian aid protects vulnerable populations, it is also easily captured by rebel groups and thus directly affects the opportunity costs of fighting. Humanitarian aid represents about 6.5% of overall aid in our sample. Here too the partial correlation of the exogenous component of predicted bilateral aid with humanitarian aid is close to zero (0.02), suggesting that our results are not driven by (unobserved) humanitarian aid.

We analyze if the effect of U.S. food aid differs from the results of overall aid presented here. Table D-8 presents the results of simple replication and modification exercises using the data from Nunn and Qian (2014). Column (1) shows that our results are qualitatively similar in the matched sample of 103 recipient countries over the period from 1975 to 2007. In column (2), we then replicate a version of their main specification, where U.S. food aid is instrumented with U.S. wheat production interacted with the probability of receiving U.S. aid. We also employ their conflict measure: a dummy for armed conflict or civil war.[12] In line with their results, we find that U.S. food aid increases the probability of continued civil conflict. Column (3) uses our fine-grained measure of conflict on the left hand side. Here we no longer observe a continuation effect for armed or civil conflicts, suggesting that their results hinge on the definition of conflict.

Last but not least, we conduct a falsification test to figure out if the identifying variation overlaps between our estimates of the effect of total ODA and the established effect of U.S. food aid. This should not be the case. Donor fractionalization of the 28 DAC donor countries should not predict U.S. food aid. Likewise, wheat production in the U.S. should not predict total ODA disbursed by the 28 DAC donors, but only a small part of U.S. overall aid. Table D-9 shows that this is reflected in the data. Hence, our primary finding that bilateral development aid promotes the continuation of small conflicts and an escalation of small to armed conflicts is quite different from the local average partial effect of U.S. food aid highlighted previously.

[11] We regress our instrument on a full set of time and country fixed effects, and obtain the residual.

[12] Note that our framework does not allow for the large set of controls used in Nunn and Qian (2014).

Table D-1

Robustness: Different linear estimation schemes

	Estimation Method:					
Variables	(1) FE-OLS	(2) CRE-CF	(3) FE-2SLS	(4) FE-OLS	(5) CRE-CF	(6) FE-2SLS
Aid to GDP (a_{2it})	-0.0011	0.0115	0.0114*	-0.0012	0.0114*	0.0116*
	(0.0011)	(0.0070)	(0.0058)	(0.0009)	(0.0062)	(0.0061)
Residuals ($\hat{\nu}_{2it}$)		-0.0127*			-0.0128*	
		(0.0073)			(0.0065)	
Interactions with Lagged States						
Small Conflict ($a_{2it} \times h_{1,i,t-1}$)				0.0058**	0.0059*	0.0077
				(0.0028)	(0.0031)	(0.0073)
Armed Conflict ($a_{2it} \times h_{2,i,t-1}$)				-0.0108	-0.0107	-0.0125
				(0.0120)	(0.0119)	(0.0162)
Civil War ($a_{2it} \times h_{3,i,t-1}$)				-0.0026	-0.0025	-0.0096
				(0.0054)	(0.0137)	(0.0104)
Lagged States						
Small Conflict ($h_{1,i,t-1}$)	0.2506***	0.2501***	0.2486***	0.2271***	0.2263***	0.2174***
	(0.0306)	(0.0302)	(0.0306)	(0.0342)	(0.0347)	(0.0439)
Armed Conflict ($h_{2,i,t-1}$)	1.1201***	1.1193***	1.1231***	1.1707***	1.1695***	1.1841***
	(0.0797)	(0.0823)	(0.0789)	(0.0996)	(0.0996)	(0.1144)
Civil War ($h_{3,i,t-1}$)	1.7902***	1.7896***	1.7899***	1.8116***	1.8105***	1.8457***
	(0.0856)	(0.0891)	(0.0835)	(0.0878)	(0.0947)	(0.1027)
Summary Statistics						
$N \times T$	4,375	4,375	4,375	4,375	4,375	4,375
T	35	35	35	35	35	35
N	125	125	125	125	125	125

Notes: All columns include the log of GDP, the log of population, recipient fixed effects, and time fixed effects. Clustered standard errors in parentheses for all columns but column (2) and (5), where we report two-step panel bootstrap standard errors in parentheses, computed with 999 replications. In columns (2), (3), (5) and (6), the Kleibergen-Paap F-statistic of the instrument is 330.8 (when the uncertainty of the bilateral stage is not accounted for) and 184.9 (when the first stage standard errors are bootstrapped by sampling over recipients in the bilateral stage). All models also estimate J cut points and the variance of the random recipient effect (not reported). * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table D-2
Alternate measures of conflict

Variables	(1) with Terror	(2) only 25 BDs	(3) UCDP-PRIO
	Perturbations on LHS		
Aid to GDP (a_{2it})	0.1302***	0.0600	0.0522
	(0.0495)	(0.0455)	(0.0465)
Residuals ($\hat{\nu}_{2it}$)	-0.1380***	-0.0722	-0.0482
	(0.0486)	(0.0451)	(0.0471)
Interactions with Lagged States			
Small Conflict ($a_{2it} \times h_{1,i,t-1}$)	0.0105	0.0200**	
	(0.0099)	(0.0085)	
Armed Conflict ($a_{2it} \times h_{2,i,t-1}$)	-0.0167	-0.0079	-0.0270
	(0.0201)	(0.0168)	(0.0202)
Civil War ($a_{2it} \times h_{3,i,t-1}$)	-0.0071		-0.0208
	(0.0248)		(0.0294)
Lagged States			
Small Conflict ($h_{1,i,t-1}$)	0.7369***	0.5280***	
	(0.0742)	(0.0806)	
Armed Conflict ($h_{2,i,t-1}$)	2.4516***	2.2636***	2.1033***
	(0.2025)	(0.1863)	(0.1870)
Civil War ($h_{3,i,t-1}$)	3.8061***		3.3420***
	(0.2456)		(0.2332)
Summary Statistics			
$N \times T$	4,375	4,375	4,375
T	35	35	35
N	125	125	125

Notes: The table shows the results of an ordered probit model with correlated random effects and a control function approach. Three-step panel bootstrap standard errors—sampling over recipients in the bilateral sample and computed with 999 replications—are in parentheses. The Kleibergen-Paap F-statistic of the instrument is 330.8 in all regressions (when the uncertainty of the bilateral stage is not accounted for) and 184.9 (when the first stage standard errors are bootstrapped by sampling over recipients in the bilateral stage). All models also estimate J cut points and the variance of the random recipient effect (not reported). * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table D-3
'Leave-one-out' test for small conflict coding

	Dependent Variable: Ordered Conflict			
	(1)	(2)	(3)	(4)
Variables	No Assassinations	No Guerrilla Warfare	No Purges	No Riots
Aid to GDP (a_{2it})	0.0774	0.0600	0.0933*	0.0630
	(0.0491)	(0.0429)	(0.0531)	(0.0454)
Residuals ($\hat{\nu}_{2it}$)	-0.0866*	-0.0688	-0.1066**	-0.0695
	(0.0483)	(0.0428)	(0.0525)	(0.0453)
Interactions with Lagged States				
Small Conflict ($a_{2it} \times h_{1,i,t-1}$)	0.0159*	0.0170**	0.0218**	0.0134
	(0.0090)	(0.0084)	(0.0091)	(0.0083)
Armed Conflict ($a_{2it} \times h_{2,i,t-1}$)	-0.0137	-0.0096	-0.0105	-0.0200
	(0.0182)	(0.0200)	(0.0189)	(0.0201)
Civil War ($a_{2it} \times h_{3,i,t-1}$)	-0.0085	-0.0046	-0.0033	-0.0125
	(0.0271)	(0.0227)	(0.0248)	(0.0294)
Lagged States				
Small Conflict ($h_{1,i,t-1}$)	0.5845***	0.3831***	0.6008***	0.7659***
	(0.0774)	(0.0759)	(0.0761)	(0.1000)
Armed Conflict ($h_{2,i,t-1}$)	2.0594***	1.9528***	2.1145***	2.1568***
	(0.1783)	(0.1689)	(0.1875)	(0.1908)
Civil War ($h_{3,i,t-1}$)	3.3906***	3.2655***	3.4314***	3.4431***
	(0.2213)	(0.2153)	(0.2214)	(0.2405)
Summary Statistics				
$N \times T$	4,375	4,375	4,375	4,375
T	35	35	35	35
N	125	125	125	125

Notes: All columns include the log of GDP, log population, the initial states, CRE at the recipient level, residual CRE, and time fixed effects. Three-step panel bootstrap standard errors—sampling over recipients in the bilateral sample and computed with 999 replications—are in parentheses. The Kleibergen-Paap F-statistic of the instrument is 330.8 in all regressions (when the uncertainty of the bilateral stage is not accounted for) and 184.9 (when the first stage standard errors are bootstrapped by sampling over recipients in the bilateral stage). All models also estimate J cut points and the variance of the random recipient effect (not reported). * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table D-4
Alternate measures of foreign aid

Variables	(1) Gross Aid over GDP	(2) with OOF	(3) No Anglo Saxon	(4) No Ang. Sax. with OOF
	Perturbations on RHS:			
Aid to GDP (a_{2it})	0.0663	0.0791*	0.0824	0.0493
	(0.0423)	(0.0474)	(0.0708)	(0.0496)
Residuals ($\hat{\nu}_{2it}$)	-0.0809*	-0.0935**	-0.1059	-0.0648
	(0.0418)	(0.0474)	(0.0683)	(0.0506)
Interactions with Lagged States				
Small Conflict ($a_{2it} \times h_{1,i,t-1}$)	0.0219**	0.0211**	0.0314**	0.0233**
	(0.0086)	(0.0092)	(0.0154)	(0.0099)
Armed Conflict ($a_{2it} \times h_{2,i,t-1}$)	-0.0119	-0.0114	-0.0208	-0.0132
	(0.0180)	(0.0190)	(0.0294)	(0.0197)
Civil War ($a_{2it} \times h_{3,i,t-1}$)	-0.0025	-0.0033	-0.0276	-0.0025
	(0.0221)	(0.0151)	(0.0546)	(0.0184)
Lagged States				
Small Conflict ($h_{1,i,t-1}$)	0.5664***	0.5730***	0.5781***	0.5780***
	(0.0779)	(0.0806)	(0.0810)	(0.0770)
Armed Conflict ($h_{2,i,t-1}$)	2.1163***	2.1132***	2.1240***	2.1238***
	(0.1842)	(0.1821)	(0.1947)	(0.1838)
Civil War ($h_{3,i,t-1}$)	3.4304***	3.4460***	3.4769***	3.4525***
	(0.2312)	(0.2162)	(0.2341)	(0.2236)
Summary Statistics				
Kleibergen-Paap F-statistic IV	264.5	148.2	95.43	169.8
$N \times T$	4,375	4,375	4,375	4,375
T	35	35	35	35
N	125	125	125	125

Notes: All columns include the log of GDP, log population, the initial states, CRE at the recipient level, residual CRE, and time fixed effects. No Anglo Saxon excludes Canada, the UK and the United States. Three-step panel bootstrap standard errors—sampling over recipients in the bilateral sample and computed with 999 replications—are in parentheses. The Kleibergen-Paap F-statistic of the instrument reported above does not account for the uncertainty of the bilateral stage. All models also estimate J cut points and the variance of the random recipient effect (not reported). * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table D-5
Adding bilateral controls

	Dependent Variable: Aid to GDP		
Variables	(1)	(2)	(3)
g_{3jt}	-0.0554***	-0.0472***	-0.0558***
	(0.0141)	(0.0144)	(0.0150)
$g_{3jt} \times \bar{p_{3ij}}$	0.2564***	0.2366***	0.2432***
	(0.0542)	(0.0593)	(0.0554)
l_{3jt}	0.3571*	0.4706	0.4370**
	(0.2099)	(0.3585)	(0.2150)
$l_{3jt} \times \bar{p}_{3ij}$	-0.7091*	-0.8614*	-0.7762**
	(0.3859)	(0.5021)	(0.3921)
Selected Controls			
UNGA Alignment	-0.0345		-0.0506
	(0.0665)		(0.0712)
Log Imports (Donor)		-0.0042***	-0.0033***
		(0.0009)	(0.0007)
Log Imports (Recipient)		0.0036***	0.0027***
		(0.0011)	(0.0008)
Additional Controls			
Dyad FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Summary Statistics			
Within-R^2	0.00432	0.00298	0.00449
$N \times \bar{T}$	104,194	122,473	100,093
$\bar{T}$	28.76	30.45	28.31
N	3623	4022	3535

Notes: The table shows the results from bilateral regressions with two-way fixed effects. Standard errors clustered on dyads are provided in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table D-6
Partialling out bilateral variables during IV prediction stage

	Dependent Variable: Ordered Conflict		
	(1)	(2)	(3)
Variables	UNGA voting	Trade	UNGA & Trade
Aid to GDP (a_{2it})	-0.0035	0.0115	-0.0330
	(0.0510)	(0.0340)	(0.0488)
Residuals ($\hat{\nu}_{2it}$)	-0.0186	-0.0266	0.0101
	(0.0494)	(0.0318)	(0.0481)
Interactions with Lagged States			
Small Conflict ($a_{2it} \times h_{1,i,t-1}$)	0.0348**	0.0205**	0.0352***
	(0.0135)	(0.0096)	(0.0136)
Armed Conflict ($a_{2it} \times h_{2,i,t-1}$)	-0.0260	-0.0092	-0.0216
	(0.0330)	(0.0202)	(0.0291)
Civil War ($a_{2it} \times h_{3,i,t-1}$)	-0.0191	-0.0018	-0.0145
	(0.0528)	(0.0236)	(0.0447)
Lagged States			
Small Conflict ($h_{1,i,t-1}$)	0.5189***	0.5651***	0.4995***
	(0.0978)	(0.0790)	(0.0918)
Armed Conflict ($h_{2,i,t-1}$)	2.2305***	2.0854***	2.1904***
	(0.2009)	(0.1819)	(0.1926)
Civil War ($h_{3,i,t-1}$)	3.6799***	3.4091***	3.6372***
	(0.2648)	(0.2170)	(0.2453)
Additional Controls			
Recipient CRE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Residual CRE	Yes	Yes	Yes
Initial States	Yes	Yes	Yes
Summary Statistics			
Kleibergen-Paap F-statistic IV	40.34	85.03	18.22
$N \times T$	3,080	4,235	3,045
T	35	35	35
N	88	121	87

Notes: The table shows the results of an ordered probit model with correlated random effects and a control function approach. The bilateral variables are partialled out of predicted aid disbursements during the IV prediction stage. UNGA voting alignment is measured as the shared 'yes' & 'no' votes between a donor and a recipient. Trade is measures by including both the log imports of donors countries for each recipient, and the log recipient imports from each donor. Three-step panel bootstrap standard errors—sampling over recipients in the bilateral sample and computed with 999 replications—are in parentheses. The Kleibergen-Paap F-statistic of the instrument reported above does not account for the uncertainty of the bilateral stage. All models also estimate J cut points and the variance of the random recipient effect (not reported). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

TABLE D-7
Additional covariates

VARIABLES	(1)	(2)	(3)	(4)	(5)
	Dependent Variable: Ordered Conflict				
Aid to GDP (a_{2it})	0.0736	0.0620	0.0448	0.0684	0.0742
	(0.0451)	(0.0450)	(0.0393)	(0.0438)	(0.0467)
Residuals ($\hat{\nu}_{2it}$)	-0.0872*	-0.0787*	-0.0611	-0.0860**	-0.0872*
	(0.0451)	(0.0432)	(0.0383)	(0.0430)	(0.0464)
Interactions with Lagged States					
Small Conflict ($a_{2it} \times h_{1,i,t-1}$)	0.0214**	0.0269**	0.0316***	0.0275**	0.0202**
	(0.0091)	(0.0109)	(0.0112)	(0.0111)	(0.0084)
Armed Conflict ($a_{2it} \times h_{2,i,t-1}$)	-0.0124	-0.0113	-0.0056	-0.0105	-0.0165
	(0.0180)	(0.0231)	(0.0183)	(0.0208)	(0.0177)
Civil War ($a_{2it} \times h_{3,i,t-1}$)	-0.0042	-0.0007	0.0018	0.0009	-0.0041
	(0.0204)	(0.0231)	(0.0157)	(0.0248)	(0.0227)
Added Controls					
Neighbor in Small Conflict	0.1246*				
	(0.0752)				
Neighbor in Armed Conflict	0.0618				
	(0.0830)				
Neighbor in Civil War	0.1652				
	(0.1083)				
Political Instability		0.2166***			
		(0.0735)			
Polity IV (revised)			-0.0132		
			(0.0085)		
Regional Polity IV				0.0184	
				(0.0178)	
GDP Growth					-1.1518***
					(0.3200)
Summary Statistics					
Kleibergen-Paap F-statistic IV	328.2	75.45	71.71	85.22	316.3
$N \times T$	4,375	3,500	3,325	3,500	4,375
T	35	35	35	35	35
N	125	100	95	100	125

Notes: The table shows the results of an ordered probit model with correlated random effects and a control function approach. Three-step panel bootstrap standard errors—sampling over recipients in the bilateral sample and computed with 999 replications—are in parentheses. All columns include the log of GDP, log population, the lagged states, the initial states, CRE at the recipient level, residual CRE, and time fixed effects. All models also estimate J cut points and the variance of the random recipient effect (not reported). * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table D-8
Our results vs. Nunn and Qian (2014)

	Dependent Variable: Ordered Conflict		
	(1)	(2)	(3)
Variables	Aid to GDP	U.S. Food aid	U.S. Food aid
Aid (a_{2it})	0.0693	0.0202	0.0131
	(0.0525)	(0.0346)	(0.0157)
Residuals ($\hat{\nu}_{2it}$)	-0.0845*	-0.0221	-0.0128
	(0.0513)	(0.0345)	(0.0157)
Interactions with Lagged States			
Small Conflict ($a_{2it} \times h_{1,i,t-1}$)	0.0306***		-0.0007
	(0.0116)		(0.0008)
Armed Conflict ($a_{2it} \times h_{2,i,t-1}$)	-0.0012	0.0031**	-0.0006
	(0.0224)	(0.0015)	(0.0008)
Civil War ($a_{2it} \times h_{3,i,t-1}$)	-0.0024		0.0001
	(0.0220)		(0.0013)
Lagged States			
Small Conflict ($h_{1,i,t-1}$)	0.5549***		0.6988***
	(0.0821)		(0.0825)
Armed Conflict ($h_{2,i,t-1}$)	2.0553***	2.0557***	2.0646***
	(0.2108)	(0.1850)	(0.1726)
Civil War ($h_{3,i,t-1}$)	3.3626***		3.2882***
	(0.2280)		(0.2271)
Summary Statistics			
Kleibergen-Paap F-statistic IV	54.16	10.45	10.45
$N \times T$	3,193	3,193	3,193
T	31	31	31
N	103	103	103

Notes: The table shows the results of an ordered probit model with correlated random effects and a control function approach. All columns include the log of GDP, log population, the initial states, CRE at the recipient level, residual CRE, and time fixed effects. Column (2) uses the classical armed conflict dummy (including civil war) as a dependent variable, otherwise our ordered measure of conflict is on the left hand side. Column (1) reports three-step panel bootstrap standard errors—sampling over recipients in the bilateral sample and computed with 999 replications—in parentheses. The Kleibergen-Paap F-statistic of the instrument reported in column (1) does not account for the uncertainty of the bilateral stage. Columns (2) and (3) report two-step panel bootstrap standard errors computed with 999 replications in parentheses. All models also estimate J cut points and the variance of the random recipient effect. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

TABLE D-9
Falsification test

	Dependent Variable:	
	(1)	(2)
VARIABLES	U.S. Food aid	Aid to GDP
Predicted aid to GDP ($\sum_j \hat{a}_{3ijt}$)	2.3173	
	(3.0231)	
Nunn and Qian (2014) IV		-0.0000
		(0.0000)
Selected Controls		
Log GDP	-53.2639	-4.7561***
	(39.6852)	(0.8251)
Log Population	78.4672*	4.2505*
	(46.1693)	(2.5013)
Additional Controls		
Country FE	Yes	Yes
Time FE	Yes	Yes
Summary Statistics		
Within-R^2	0.0439	0.115
$N \times T$	3,193	3,193
T	31	31
N	103	103

Notes: The table shows the results from two-way linear fixed effects models. Standard errors clustered on recipients are provided in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Additional references

Angrist, J. D. and J.-S. Pischke (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press.

Bailey, M. A., A. Strezhnev, and E. Voeten (2017). Estimating dynamic state preferences from United Nations voting data. *Journal of Conflict Resolution 61*(2), 430–456.

Banks, A. S. and K. A. Wilson (2015). *Cross-National Time-Series Data Archive*. Jerusalem: Databanks International.

Bosker, M. and J. de Ree (2014). Ethnicity and the spread of civil war. *Journal of Development Economics 108*, 206–221.

Campos, N. and M. Gassebner (2013). International terrorism, domestic political instability and the escalation effect. *Economics & Politics 25*(1), 27–47.

Kuziemko, I. and E. Werker (2006). How much is a seat on the Security Council worth? Foreign aid and bribery at the United Nations. *Journal of Political Economy 114*(5), 905–930.

Nunn, N. and N. Qian (2014). U.S. food aid and civil conflict. *American Economic Review 104*(6), 1630–1666.

START (National Consortium for the Study of Terrorism and Responses to Terrorism) (2016). Global terrorism database [Data file]. https://www.start.umd.edu/gtd.

Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data* (2nd ed.). Cambridge (MA): MIT Press.