Foreign Aid Shocks as a Cause of Violent Armed Conflict*
Supporting Information Appendix

Richard A. Nielsen, Michael G. Findley, Zachary S. Davis,
Tara Candland, Daniel L. Nielson

September 17, 2010

This document offers more detailed discussion of the analysis conducted for the article and presents the results of numerous statistical models that we examined but, due to space constraints, were not able to present in the article. The results of the additional analysis offer substantial support for the argument advanced in the article and the tests make clear the conditions under which the relationship between *aid shocks* and conflict applies.

The additional analysis covers three areas: the results using alternative measures of our key independent and dependent variables, issues of estimation and endogeneity, and miscellaneous tests to probe the robustness of results. We also include a list of variables and their operational definitions at the end of this document.

Additional Specifications of Key Variables

A1. Using OECD Aid Data Alone

By adding more than 40 donors and nearly $2 trillion to tracked development finance flows, AidData is a clear improvement over the existing OECD aid information. But because the OECD data on aid has been the standard in the literature, we re-estimated our models using only OECD data to analyze whether our findings are consistent across both datasets.

To make an appropriate comparison, we first describe and compare aid flows as reported by AidData to those reported by the OECD. Figure A1 shows the comparison for four selected aid recipients. We note that although the two measures of aid generally track well, they diverge radically in some specific instances. In particular, we note that the OECD totals are generally lower than the AidData numbers, providing evidence that the additional aid contained in AidData is significantly greater than the OECD. We also note a few cases in which OECD figures are larger than AidData figures. This may occur because AidData takes greater care than the OECD to eliminate duplicate projects from the data.

* We thank Robert Bates, Andrew Coe, David Davis, Paul Diehl, Matthew Kocher, Jim Kuklinski, Rebecca Nielsen, Robert Powell, Carie Steele, Beth Simmons, Michael Tierney, Dustin Tingley, Brian Urlacher, members of the Project Level Aid Database (PLAID) research team at Brigham Young University, participants at the 2007 annual meetings of the International Studies Association, and five anonymous reviewers at the AJPS for helpful comments and advice. Richard Nielsen is supported by a National Science Foundation Graduate Research Fellowship. AidData (formerly known as Project-Level Aid or PLAID) was supported by the Bill & Melinda Gates Foundation; the William and Flora Hewlett Foundation; National Science Foundation grant SES-0454384; the College of William and Mary; and the College of Family, Home and Social Sciences, the Department of Political Science, and the David M. Kennedy Center for International Studies at Brigham Young University. Replication materials and this Supplemental Information appendix are available at http://dvn.iq.harvard.edu/dvn/dv/rnielsen and http://politicalscience.byu.edu/faculty/mfindley/.

These discrepancies are not trivial. Some cases in the OECD data appear to indicate an aid shock, whereas AidData shows that the country-year passed without a shock. Comparing the number of overlapping aid shocks makes these discrepancies stand out: only 246 of the 393 aid shocks in AidData and the 330 aid shocks in the OECD databases are the same, meaning that many aid shocks are not common to both datasets.

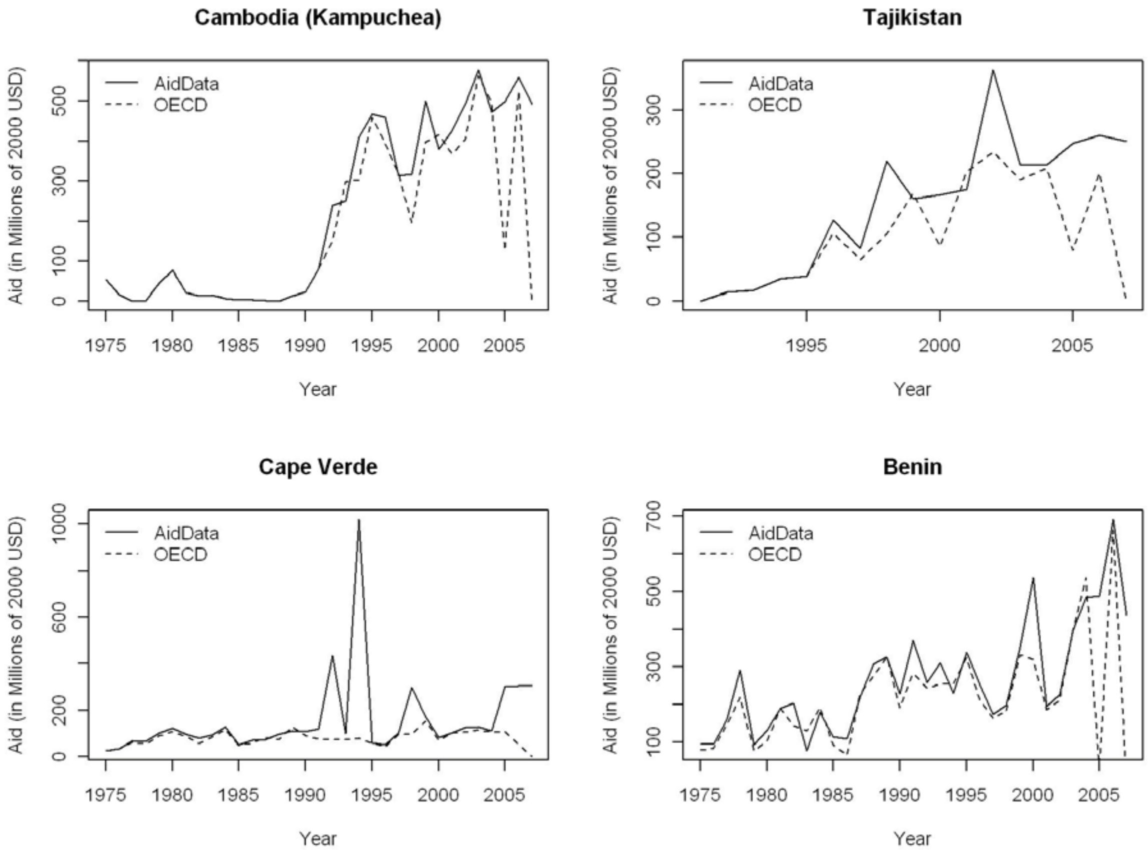

Figure A1: *Comparison of aid using AidData and OECD data in four randomly selected countries.*

Table A1 shows the results of our primary specifications (rare events logit, fixed effects logit, propensity score matching, and genetic matching) using only the OECD aid data. The results indicate that, while similar, our findings would have been less robust across alternative specifications if we had used the less complete OECD data. In particular, the results are weaker in the fixed-effects logit and with propensity score matching when we use the OECD data. It is encouraging that our most rigorous model – genetic matching – gives a result consistent with the results we obtained using AidData information. Still, we note that if we employed only the OECD data in our analysis, our conclusions would be on shakier ground statistically.

The difference between these results and those presented in the paper nicely points out the realities of measurement error. Because we have the AidData figures, we can obtain a rough estimate of how far off the OECD data are from what we might learn from more comprehensive data (AidData is not without problems, but it does capture the universe of development finance more fully). The weakened findings we obtain in some models when using the OECD measure are likely due to measurement error that attenuates the estimated causal effects of aid shocks. Because we have substantial evidence that the AidData database captures more aid transfers than the OECD database, we can more confidently stand by the results reported in the main paper and discount the findings using the OECD data.

Table A1: Models with OECD aid data

	(1) Rare Events Logit	(2) Fixed Effects Logit	(3) Propensity Score Matching	(4) Genetic Matching
Aid Shock	0.744**	0.578*	0.556	0.929**
	(0.326)	(0.350)	(0.395)	(0.376)
Positive Aid Shock	-0.0401	-0.295		
	(0.384)	(0.398)		
Human Rights Violations	0.609***	0.539***	0.502**	0.505**
	(0.140)	(0.205)	(0.215)	(0.216)
Assassinations	0.140	0.185	0.270	0.117
	(0.0980)	(0.135)	(0.245)	(0.187)
Riots	0.0117	-0.157	0.182	0.173
	(0.134)	(0.163)	(0.160)	(0.170)
General Strikes	-0.00488	0.555*	0.379	0.398
	(0.206)	(0.300)	(0.266)	(0.257)
Anti-Gov. Demonstrations	-0.0460	-0.104	-0.216	-0.0718
	(0.123)	(0.122)	(0.258)	(0.220)
Infant Mortality	0.00319	-0.0177	0.00356	0.000703
	(0.00472)	(0.0198)	(0.00638)	(0.00753)
Bad Neighborhood	-0.0360	-0.0523	-0.0365	0.0545
	(0.117)	(0.203)	(0.194)	(0.185)
Partial Autocracy	0.261	0.279	0.485	0.268
	(0.335)	(0.560)	(0.469)	(0.481)
Partial Democracy	-0.648	-0.0569	0.228	-0.434
	(0.470)	(0.800)	(0.666)	(0.673)
Factional Democracy	0.688*	1.519**	0.836	0.399
	(0.389)	(0.650)	(0.586)	(0.635)
Full Democracy	0.186	1.073	1.636	0.722
	(0.553)	(1.194)	(1.141)	(0.969)
ln(GDP per capita)	-0.231	-0.589	-0.294	-0.557
	(0.232)	(0.766)	(0.282)	(0.398)
ln(Population)	0.0634	-2.609	-0.0763	0.000721
	(0.0867)	(1.905)	(0.191)	(0.246)
Oil	0.0101***	0.00622	0.115	0.0372
	(0.00303)	(0.00861)	(0.129)	(0.243)
Instability	0.254	0.0420	0.298	0.348
	(0.262)	(0.375)	(0.406)	(0.439)
Ethnic Frac.	1.355**		0.626	-0.389
	(0.581)		(1.005)	(1.100)
Religious Frac.	-0.762		0.801	0.0238
	(0.677)		(1.467)	(1.222)
Non-contiguous	0.994***		-0.736	0.602
	(0.324)		(0.867)	(0.782)
Mountains	0.0878		0.0499	-0.111
	(0.0931)		(0.135)	(0.152)
Cold War	0.120	-0.664	-0.780	-0.266
	(0.283)	(0.548)	(0.520)	(0.443)
Constant	-5.040*		-1.808	-0.0344
	(2.882)		(3.608)	(4.365)
Observations	2627	953	786	698

*Models with OECD aid data. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1*

A2. Using Aid Disbursement Data for 2001-2005

In the main specifications appearing in the article, we used data on aid commitments rather than aid disbursements because no reliable data on disbursements exist prior to roughly 2002. Indeed, the OECD's online user guide notes, "[a]s to the analysis on CRS disbursements it is not recommended for flows before 2002, because the annual coverage is below 60%, while it is around and over 90% since 2002 and reached nearly 100% starting with 2007 flows" (OECD 2010). Our analysis suggests that disbursement data before 2001 has significantly less coverage than sixty percent. And, of course, because of data availability on covariates, our analysis of disbursements can only run through 2005, which truncates the available sample significantly.

Nevertheless, to explore what might be possible with disbursement data, we used disbursements reported by the OECD in addition to that collected by the AidData project. Over time, as we note in the article, commitments and disbursements are highly correlated. There may, however, be differences in the short run, particularly given that development banks disburse loans in "tranches" over a multi-year period. Moreover, missing disbursement data causes even greater divergence between commitments and disbursements.

We show that disbursement amounts differ markedly from the commitment data in the levels of aid captured by plotting commitments and disbursements for four randomly selected countries, shown in Figure A2. Although the levels of commitments vs. disbursements are quite different, we note that the trends in the graphs share broad contours, which is generally consistent with findings by others that commitments and disbursements are closely correlated (Neumayer 2003c, 43; Nielson and Tierney 2005, 789; also reported in the article). The difference in levels is likely due to missing disbursement data.

These figures suggest possible problems with using disbursement data: they offer a much more limited picture of aid flows. Some of these differences may be genuine – aid that is committed in a single year is often disbursed over multiple subsequent years, or donor priorities may change between the time of commitment and disbursement. However, it seems unlikely that this can account for the differences in levels we observe in the data. Instead, it seems probable that disbursement data are dramatically under-reported, even after 2002 when these figures are supposedly more reliable. In particular, we note that some large decreases that appear in the commitment data are not reflected in the disbursement data.

Sorting out the disconnect between commitment and disbursement data are beyond the scope of this paper. But we nonetheless offer an initial test of our argument using the most reliable disbursement data and constraining our analysis to the years 2001-2005.

We define a disbursement aid shock using the same aid over GDP cutoff we used for defining an aid shock with commitment data (*Aid Change/GDP* < - 0.0054). This means that our disbursement shocks have the same magnitude as our commitment shocks. We then estimate the main logistic regression specification from the paper using *Disbursement Aid Shock* as the key predictor and including only observations between 2001 and 2005 (when the OECD claims the disbursement data are most reliable).

In this model, we find that *Disbursement Aid Shocks* have a positive effect on the probability of conflict onset, but that the effect is only significant at the .1 level ($b = 2.63$, $p = 0.077$). We find it remarkable that the result approximates statistical significance – there are merely four conflict onsets left in the restricted dataset, only one of which is preceded by a disbursement shock.

These data seem so problematic and the model hinges on so few observations of conflict that we hesitate to draw any strong conclusions. Still, it is heartening that the results manage to offer some support for the argument advanced in the paper.

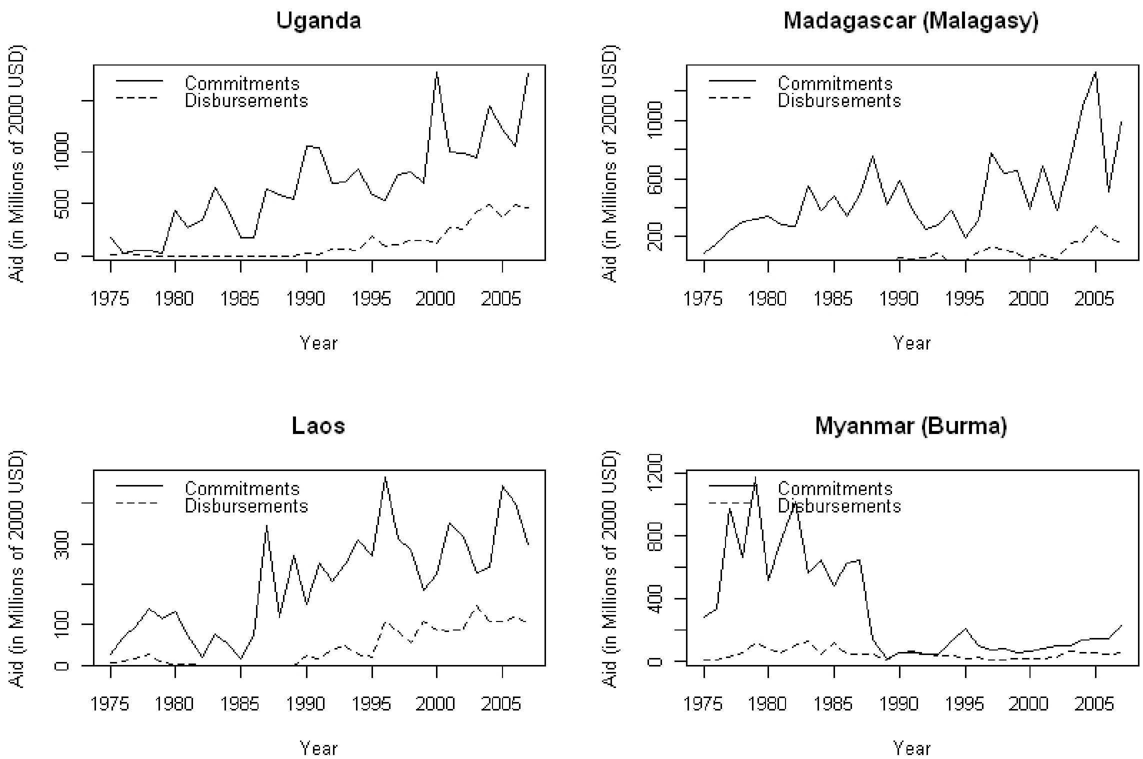

Figure A1: A comparison of aid flows measured using commitment data (solid line) and disbursement data (dashed line) for randomly selected countries.

A3. Alternative Measures of Aid Shocks

In addition to considering data source and commitments vs. disbursement, we next consider some alternative specifications for defining aid shocks. It is plausible that sustained aid shocks may be more destabilizing than isolated shocks. To address this, we average one-year aid shocks for the five-year period previous to each observation, thus measuring the number of the years out of the last five that a country experienced an aid shock. Here, we find that the average number of aid shocks in the last five years has a strong effect on the probability of conflict onset ($b = 1.96, p = 0.013$).

By this logic, the less sustained the aid shock, the less likely it should be to affect armed conflict onset. To consider this, we also tested a measure of aid shocks that was based on the amount of aid lost only in the previous year (rather than in the previous two years). With this measure of aid shocks, our key finding for aid shocks is only significant at the .1 level ($b = 0.50$, $p = 0.079$). This suggests that sustained shifts in aid over two or more years are more destabilizing than one-year shifts.

A4. Military Aid

We expect that our hypotheses apply to military aid as well as economic aid. Unfortunately, we are unable to test this fully because most donors do not report the amount of military aid they provide. The United States does report its aid, however, so we estimated a series of models in which we combined economic aid with US military aid (USAID 2010), both standardized in USD 2000, and then recalculated the *Aid Shocks* variable and used this new variable in the main models we present in the text.

We find that the effect of aid shocks grows slightly stronger when we include US military aid. In the rare-events logit, we find that the coefficient on the combined economic-military *Aid Shock* variable is 0.95 ($p = 0.001$). The models with fixed effects ($b = 0.99$, $p = 0.005$), propensity score matching ($b = 0.74$, $p = 0.052$), and genetic matching (b =,0.99 $p = 0.009$) also suggest a significant effect for *Aid Shocks*.

A5. Measuring Aid Changes Continuously Rather than as Shocks

In addition to considering alternative definitions and data to define aid shocks, we also test our models using continuous changes in aid as the key causal variable of interest. As discussed in the paper, continuous changes in aid do not capture well shifts in the distribution of power identified in the credible commitment logic (Powell 2004) and specified in our paper. Thus using a continuous measure helps us test whether the argument we advance in the paper about extreme negative changes is at work or whether any changes – be they small or large – affect the probability of armed conflict onset.

We measure the lagged change in aid as a percentage of GDP, averaging change in the previous two years to account for the time gap between aid commitments and the time at which countries actually receive (or do not receive) the aid. We expect the direction of the *Aid Change* variable to be negative: as negative changes become smaller and positive changes become larger, the likelihood of armed conflict onset should decrease.

We find that the variable *Aid Change* is negatively correlated with conflict – meaning that negative aid changes are more likely to lead to violence – but this effect is not statistically significant (b = -10.1, p = 0.24). When we examine *only* negative aid changes as our theory suggests, we again find a negative sign that is statistically insignificant (b = -15.0, p = 0.139). From these models, we conclude that the relationship between aid changes and conflict onset is not linear. This supports our argument that rapid shifts in the distribution of power occurring from extreme negative changes in aid have categorically different effects on violence than small changes in aid flows or positive changes.

A6. Positive Aid Shocks at Different Thresholds

As discussed in the article, we also consider the possibility that positive aid shocks incite civil conflict by empowering the government. We operationalize positive aid shocks analogously to negative aid shocks, defining changes in aid that were above the 85th percentile as *Positive Aid Shocks*. The results shown in the main body of the paper indicate that positive aid shocks have no appreciable impact on conflict onset. However, the effect of positive aid shocks might be sensitive to the arbitrary threshold used to define a shock. Here, we vary the cutoff, setting it as low as the 65th percentile and as high as the 95th percentile to observe how the coefficient of *Positive Aid Shocks* in the main specification changes as we alter the threshold for this variable. In particular, we are interested to understand whether there are any thresholds that would generate a statistically significant result for *Positive Aid Shocks*. As can be seen in Figure A3, the confidence bands encompass zero for all thresholds of *Positive Aid Shocks*, suggesting that positive shocks do not provoke civil conflict in a significant way at any level.

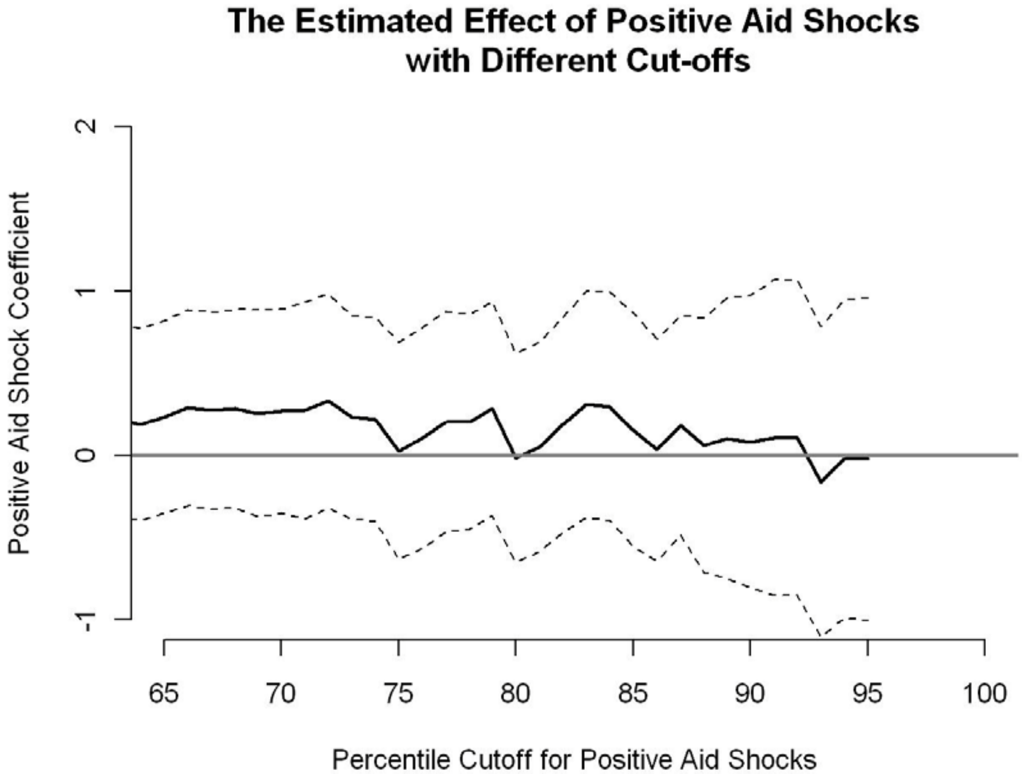

Figure A2: *The coefficient and 95 percent confidence bands on the variable* Positive Aid Shock *at different percentile cutoffs used to define a positive aid shock.*

A7. Alternative Methods of Operationalizing Conflict Onset

In our primary specifications, we drop subsequent years of conflict from the dataset, but code resumed conflict (after at least a year of peace) as a new onset. In Table A2, we show the results of models with different definitions of conflict onset. Column 1 shows the results when we included conflict years (after the onset year) in the dataset with *Conflict Onset* coded as zero.

In column 2, we only include the first conflict onset after 1980 for any given country, dropping all observations after the first onset. The logic behind this alternative is to be agnostic about whether later occurrences of armed conflict are related to the first instance (Sambanis 2004), thus avoiding the sticky issue of deciding whether flare-ups in countries with conflicts are simply continuations of the previous conflict or are new conflicts altogether.

In column 3, we only include purely internal conflicts ("type 3" conflicts in the UCDP/PRIO Armed Conflict Dataset), rather than including both internal and "internationalized internal" ("type 4") conflicts. These first three variations are all consistent with the main results reported in the paper. In column 4, we include only the 46 conflict onsets in the sample that eventually reached 1,000 deaths or more. This is the only result that is not fully consistent with the primary specification: we find that aid shocks are only statistically significant at the .1 level, suggesting that aid shocks are more predictive of small conflicts than large conflicts.

Table A2: Alternative Operationalizations of Conflict Onset

	(1) Subsequent conflict-years included	(2) Only the first onset included	(3) "Internationalized Internal" conflicts excluded	(4) Only onsets that eventually reach 1000 deaths
Aid Shock	0.736***	1.135***	0.840***	0.658*
	(0.251)	(0.430)	(0.296)	(0.362)
Positive Aid Shock	0.0116	0.333	0.143	-0.904
	(0.357)	(0.462)	(0.396)	(0.729)
Human Rights Violations	0.0694	0.427*	0.610***	0.661***
	(0.108)	(0.231)	(0.157)	(0.228)
Assassinations	0.0694	0.147	0.152	0.140
	(0.0696)	(0.173)	(0.0971)	(0.146)
Riots	0.00226	0.204	0.0114	0.167
	(0.0688)	(0.155)	(0.132)	(0.120)
General Strikes	0.0322	0.0565	0.0549	-0.414
	(0.181)	(0.247)	(0.213)	(0.441)
Anti-Gov. Demonstrations	0.0357	-0.289	-0.0235	-0.222
	(0.0657)	(0.187)	(0.122)	(0.144)
Infant Mortality	0.00467	0.00483	0.000618	0.00172
	(0.00456)	(0.00619)	(0.00519)	(0.00904)
Bad Neighborhood	0.0312	-0.0778	-0.0339	0.00456
	(0.109)	(0.140)	(0.121)	(0.184)
Partial Autocracy	0.262	0.159	0.366	0.378
	(0.313)	(0.456)	(0.327)	(0.445)
Partial Democracy	-0.380	-0.843	-0.521	-0.550
	(0.445)	(0.699)	(0.468)	(0.698)
Factional Democracy	0.650*	0.792	0.617	0.612
	(0.354)	(0.492)	(0.389)	(0.573)
Full Democracy	-0.0913	-0.277	0.0778	0.0606
	(0.552)	(1.003)	(0.619)	(0.698)
ln(GDP per capita)	-0.154	-0.181	-0.228	-0.382
	(0.217)	(0.331)	(0.256)	(0.403)
ln(Population)	0.0553	0.143	0.131	0.130
	(0.0824)	(0.130)	(0.0883)	(0.139)
Oil	0.00956***	0.00825**	0.00734	0.0133***
	(0.00252)	(0.00364)	(0.00567)	(0.00450)
Instability	0.129	0.392	0.297	0.165
	(0.250)	(0.405)	(0.285)	(0.391)
Ethnic Frac.	0.911*	1.126	1.574**	1.361
	(0.545)	(0.699)	(0.651)	(1.012)
Religious Frac.	-0.431	-0.0930	-0.881	-1.047
	(0.594)	(0.840)	(0.779)	(1.337)
Non-contiguous	0.472	1.263**	0.891**	0.967**
	(0.295)	(0.595)	(0.416)	(0.491)
Mountains	0.0401	0.197	0.0661	0.101
	(0.0802)	(0.125)	(0.0972)	(0.187)
Cold War	0.153	-0.0257	-0.0285	0.279
	(0.295)	(0.483)	(0.269)	(0.392)
Constant	-4.464	-7.201*	-6.082**	-5.318
	(2.819)	(4.136)	(3.010)	(5.136)
Observations	3142	2061	2627	2627

*Rare events logistic regression. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1*

Estimation and Endogeneity

A8. Improvement in Covariate Balance through Matching

Statistical theory shows that a matched sample with similar propensity scores will generally have similar distributions of each of the individual covariates (Rosenbaum and Rubin 1983); we confirm this in Figure A4 by showing the improvement in balance obtained in the matched datasets, both for propensity-score matching and for genetic matching.

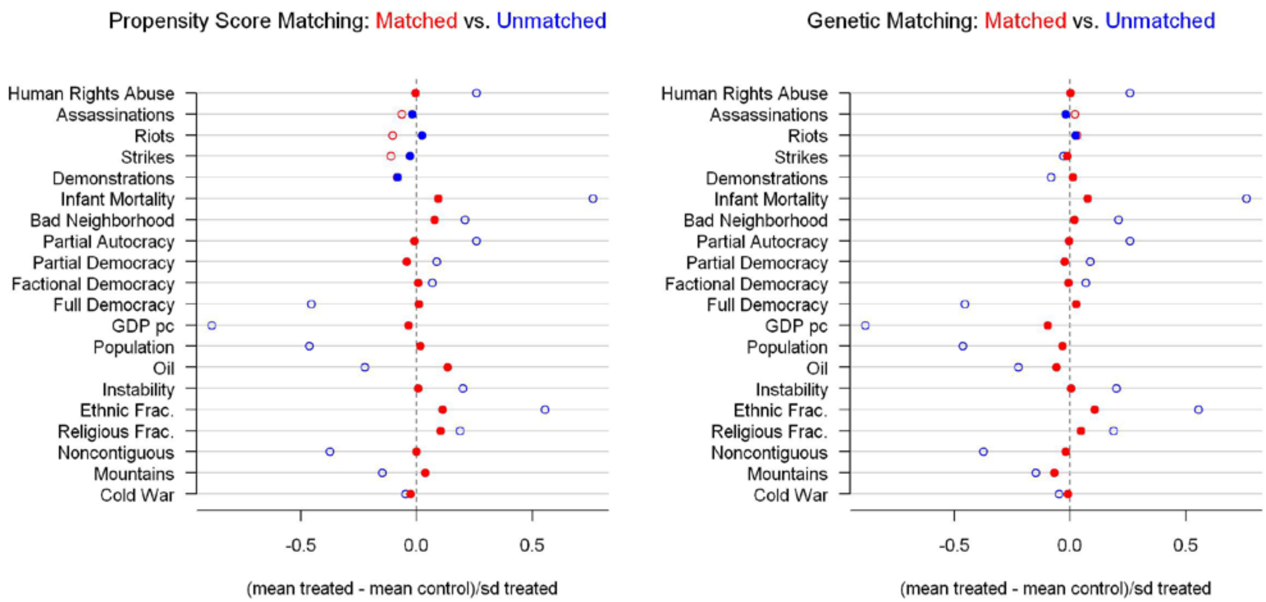

Figure A4: *The improvement in mean balance obtained from propensity score matching and genetic matching. For each variable and dataset (matched or unmatched), we subtract the mean of the controls from the mean of the treated units and standardize each by that variable's standard deviation in the treated group. Values close to zero indicate better mean balance. Filled circles are closer to zero than corresponding open circles.*

A9. Instrumental Variables

We considered a number of possible instrumental variables (IV) to account for endogeneity, but ultimately we abandoned the IV strategy because either the instruments were not significantly correlated with aid shocks or they violated the crucial exclusion restriction.[1] Thus, we rejected even the instrument – membership on the United Nations Security Council – that offered some statistical support for the argument advanced in the paper. Although instruments have been common in the economics literature and are increasingly used in political science, compelling arguments from a number of sources including Deaton (2009), Easterly (2009), and others indicate that instruments need to be chosen with extreme caution to be able to meet the stringent conditions for their effective use (e.g. Heckman 1995; Stock, Wright, and Yogo 2002; Murray 2006). What is more, many years of

[1] We were also unable to use Hausman-style tests of endogeneity that rely on having valid instruments.

research using instruments in the aid literature has accomplished less than hoped in sorting out endogeneous relationships. Thus, we proceed cautiously, considering three potential instruments along with their strengths and weaknesses.

For a variable to be a valid instrument, two conditions have to hold: (1) it needs to be significantly correlated with the selection variable and (2) it must meet the exclusion restriction, where it cannot be related to the ultimate dependent variable except through the selection variable. Thus, in our case, the potential instrument must have a strong partial correlation with *Aid Shocks* in the presence of the other exogenous regressors (the IV terminology for the other control variables in our model of conflict onset). Additionally, it must be as good as randomly assigned – that is, it can have no correlation with any variables or set of variables that influence conflict onset except through aid shocks. Note that this precludes any instrument that has a direct effect on conflict that is not mediated by aid shocks, as well as any instrument that is potentially correlated with (either caused by or causing) the underlying determinants of conflict. We judge each potential instrument against these criteria.

United Nations Security Council rotating membership

Work by Kuziemko and Werker (2006) shows that as recipient countries are elected to serve as one of ten rotating members of the UN Security Council, they receive substantially more foreign aid. It is plausible, then, that these same countries receive substantially less foreign aid once they rotate off of the Security Council, so we considered this as a potential instrument for *Aid Shocks*. The UNSC is composed of five permanent members and ten rotating members that serve for two years each. These ten rotating seats are elected by secret ballot, with three designated seats for African states, and two designated seats each for the regions of Latin America and the Caribbean, Asia, and Western Europe and Others. UNSC members receive significantly more foreign aid during their tenure on the council, presumably because of their increased influence in the UN during this time.

To use this potential instrument, we would need to show that UNSC membership also predicts aid shocks – in particular, that UNSC members would be substantially *less* likely to experience an aid shock than non-UNSC members. We find that after we condition on other covariates, the relationship between UNSC membership and *Aid Shocks* is not significant statistically (the F-statistic is roughly 1). And weak instruments can be severely problematic even if they satisfy the exclusion restriction.

Although this instrument might plausibly meet the exclusion restriction, we are somewhat skeptical. In particular, we are concerned that membership on the Security Council has a direct dampening effect on the probability of civil conflict or that UNSC members are "different" in some way that makes them less likely to experience rebel violence. The process by which UNSC members are chosen is fairly secretive (including a secret ballot), so there is no regression discontinuity design to leverage here.

Despite reasons not to use the UNSC membership instrument, we still estimated an instrumental variables regression and find strong support for a relationship between aid shocks and the onset of violent armed conflict. As discussed, however, we are skeptical about relying too heavily on this apparently supportive result because the instrument is weak and likely invalid.

United Nations "Friends"

We also considered voting similarity between recipients and donors as a potential instrument. Voting similarity in the UN might be considered a measure of the shared interests between donors and recipients and is included in some studies of aid allocation. Specifically, we use the Affinity data (Gartzke 1998) and define a recipient-year to be a "friend" of the OECD donors when it is in the top quartile of states with the closest similarity to the US votes in the UN.[2] In the first stage regression, the instrument *UN Friend* is a significant negative predictor of aid shocks.

However, we think it even less likely that the exclusion restriction holds for this instrument. In particular, we expect that countries that are "friends" with the US are likely to receive assistance other than foreign aid – including security alliances, troop support, arms shipments, and favorable trading conditions, among others – that influences the probability of civil conflict outbreak. If this direct effect exists, then this instrument is invalid.

In the instrumented model, we find that *Aid Shocks* has no significant effect on conflict onset. The estimated effect is positive but the standard errors are very large. Again, because we strongly doubt that this instrument is valid, we do not place significant stock in this finding.

Changes in Donor GDP

We also consider "changes in average donor GDP" from de Ree and Nilleson (2009). It is intuitive that aid shocks might be more likely as donor GDP decreases: donors looking to tighten their belts are likely to consider cutting off aid flows. To establish the relevance of this potential instrument, we calculate changes in donor GDP by summing the real GDP per capita of the OECD donors in each year. We then take the first difference, subtracting the total donor GDP in year *t*-1 from the total donor GDP in year *t*.

We expect that as changes in total donor GDP become more positive, aid shocks will be significantly less likely. Conversely, as changes in total donor GDP become more negative, aid shocks will be increasingly likely. Indeed, we find that this measure is correlated at the .1 level with aid shocks in the first stage regression, so the instrument is not as strong one we might like. However, there are also compelling reasons to doubt that changes in donor GDP meet the exclusion restriction necessary to be a valid instrument, which we discuss below.

De Ree and Nilleson argue that donor GDP is a valid instrument donors often have goals to provide a certain portion (say, 0.7 percent) of their GDP as aid, so donor GDP will be correlated with aid flows while remaining uncorrelated to conflict (or the propensity for conflict) in recipient countries. There are several problems with this line of argument that together make this instrument unsuitable. First, donor GDP may be correlated with the health of recipient economies through global business cycles. If recipients with poor economies are more prone to conflict, the exclusion restriction is likely to be violated because the instrument is not assigned "as if" at random. That is, if developed economies are experiencing hard times, it stands to reason that the developing economies that depend on wealthier countries for trade, financial flows, and investment may also be struggling – and the economic downturn may provoke conflict (Miguel, Satyanath, and Sergenti 2004). Second, if donors withdraw aid from their client governments, they are likely to withdraw other types of support as well (e.g., the promise of military assistance to help put down rebellion). Thus, fluctuations in donor GDPs might have a strong direct effect on conflict that does not flow solely through *Aid Shocks*.

So we take exception to the claims by de Ree and Nillesen that "U.S. GDP is unlikely to be systematically related to specific characteristics at the recipient-country level and is therefore a good candidate for a proper instrument" (310). In fact, the authors effectively admit to the validity of this objection when they note that it is "essential" to include a number of covariates to make the exclusion restriction plausible (311). Unfortunately, once we must include covariates to make a restriction

[2] We use US votes as a proxy for the interests of OECD donors generally. One could think of instead calculating a weighted affinity score, where the weights are determined by the strength of aid ties between donors.

exclusion hold, we are once again subject to the same problems of omitted variables that we face when matching, and, unlike matching, it is much more difficult to tell how well we are succeeding at creating similarity between our treated and control groups.

If donor GDP is a valid instrument for aid shocks, then the results with the instrumented model are very surprising: we find that aid shocks dramatically *decrease* the probability of conflict onset. This in itself would be a substantial finding, and quite counter-intuitive. However, we think it much more likely, given its highly probable violation of the exclusion restriction, that the instrument is simply not valid.

We should also note that we discarded a number of other potential instruments on similar grounds without fully specifying them econometrically. Because we are introducing a new variable – aid shocks – to the literature on foreign aid, we have very little guidance about what factors cause aid shocks. Still, we combed the aid allocation literature for other variables that might affect aid flows while being plausibly exogenous to civil conflict. We considered trade ties, alliances, and other measures of affinity with aid donors. None of these variables seemed remotely exogenous to us, so we excluded them at the outset.

Overall, we find that very little can be concluded from instrumental variables for a simple reason: we could find no credible instruments.

Table A3: Possible Instruments

Potential Instrument	**Strong predictor of aid shocks?**	**Justifiable exclusion restriction?**	**Results**
UN Security Council Rotating Membership	No. $b = -0.02$, $p = 0.31$	Probably not. (1) States that are likely to be elected to the UNSC are probably less likely overall to have conflict.	Aid Shocks *increase* conflict onset. $b = 3.01$, $p < 0.0001$
UN Friend	Yes. $b = -.047$, $p = 0.012$	No. (1) Being "UN friends" with the US probably means that the US (and other OECD donors) are more likely to lend military assistance (other than aid) that deters rebels.	Aid Shocks have a positive but insignificant effect. $b = 0.64$, $p = 0.85$
Changes in Average Donor GDP	Marginally (at 10 percent level). $b = -0.00000023$ $p = 0.051$	No. (1) Donor GDPs may be correlated with the health of recipient economies through global business cycles, and recipients with poor economies are more prone to conflict. (correlated with underlying causes) (2) If donors withdraw aid from their client governments during hard times, they will also with withdraw support of other types. (direct effect)	Aid shocks *decrease* conflict onset. $b = -2.42$, $p < 0.0001$

This table lists information about the potential instruments attempted, including our assessment of the plausibility of the assumptions necessary for the instrument to be valid.

A10. Does Future Violence "Predict" Past Aid Shocks

We have discussed the endogeneity problem that would arise if aid shocks occur in anticipation of conflict onset. In addition to the solutions we propose, we also show evidence that future violence is not predictive of a past aid shock. Specifically, we estimate several models[3] that include the two-year *leads* of conflict onset to predict aid shocks and find that future conflict does not predict past aid shocks well. To be sure, this test has its own problems – if aid shocks cause future conflict, then this correlation could still exist with the regression equation reversed. Still, it is encouraging that future conflict is not a significant predictor of past aid shocks using both the one-year lead of conflict (b = -.42, p = 0.36) and the two-year lead of conflict (b = -.26, p = 0.44) in a logistic model predicting *Aid Shocks* with country random effects (results are similar using fixed effects). This gives us more confidence that donors are not accurately anticipating conflict onset and withdrawing aid wherever conflict is likely.

Additional Robustness Checks

A11. Influential Observation

To check model fit and ensure that the results are not simply the result of a few influential observations, we calculate studentized residuals, deviance residuals, and leverage and plot these against predicted values to identify which observations are most influential in the model. We then re-estimate the primary specification excluding the most influential observations (by each measure of influence). The results have the same substantive interpretation as those presented in the article, and in most models, excluding the most influential observations *increases* the estimated effect of aid shocks.

The specific models are shown in Table A4. In the first column of this table, we omit 24 observations that have studentized residuals larger than 5. In the second column, we calculate deviance residuals and omit 73 observations that have deviance residuals large than 2. The third column omits the 57 observations that have the highest leverage in the model (observations with "hat" values greater than 0.05). The results of all three models indicate that our findings are not the result of a few influential outlying cases.

[3] We use logistic regression with fixed or random country intercepts, and alternative including or excluding the lags of conflict onset in addition to the leads of conflict onset.

Table A4: Excluding Influential Observations

	(1) Excluding obs. with studentized residuals > 5	(2) Excluding obs. with deviance residuals > 2	(3) Excluding obs. with hat values > 0.05
Aid Shock	1.331***	4.267**	0.811***
	(0.358)	(1.940)	(0.281)
Positive Aid Shock	0.0580	1.814	-0.0737
	(0.477)	(1.551)	(0.396)
Human Rights Violations	0.912***	2.298***	0.501***
	(0.175)	(0.740)	(0.170)
Assassinations	0.161	0.313	0.213
	(0.124)	(0.238)	(0.140)
Riots	-0.0618	0.175	0.224
	(0.153)	(0.209)	(0.160)
General Strikes	0.123	0.744	0.121
	(0.231)	(0.679)	(0.246)
Anti-Gov. Demonstrations	0.000188	-0.366	-0.211
	(0.127)	(0.223)	(0.133)
Infant Mortality	0.00133	0.0138	0.00283
	(0.00583)	(0.0127)	(0.00498)
Bad Neighborhood	0.0833	0.0799	-0.0727
	(0.133)	(0.210)	(0.121)
Partial Autocracy	0.311	2.342**	0.351
	(0.375)	(1.165)	(0.374)
Partial Democracy	-0.987*		-0.650
	(0.595)		(0.522)
Factional Democracy	1.043**	3.092**	0.683
	(0.431)	(1.277)	(0.468)
Full Democracy	0.558		0.404
	(0.614)		(0.587)
ln(GDP per capita)	-0.170	0.245	-0.269
	(0.296)	(0.866)	(0.231)
ln(Population)	0.0482	0.354	0.0689
	(0.0978)	(0.276)	(0.0915)
Oil	0.0121***	0.0420***	0.0131***
	(0.00345)	(0.0114)	(0.00315)
Instability	0.298	1.017	0.0320
	(0.308)	(0.637)	(0.299)
Ethnic Frac.	2.429***	5.104*	1.722***
	(0.664)	(2.833)	(0.593)
Religious Frac.	-0.883	0.588	-0.962
	(0.867)	(2.531)	(0.666)
Non-contiguous	1.259***	1.130	0.820***
	(0.392)	(0.938)	(0.306)
Mountains	0.0360	0.692	0.0865
	(0.121)	(0.437)	(0.0973)
Cold War	0.239	0.971	0.148
	(0.291)	(0.596)	(0.344)
Constant	-7.261**	-30.47**	-4.690
	(3.639)	(15.24)	(2.895)
Observations	2603	2554	2570

*Robust standard errors in parentheses. Standard errors clustered by country. Cubic splines included but not shown. *** p<0.01, ** p<0.05, * p<0.1. The variables Partial Democracy and Full Democracy are excluded from model 2 because they perfectly predicted one of the outcomes after 73 influential observations were removed.*

A12. The (Non-)Effects of Aid Dependence on Vulnerability to Aid Shocks

While influential aid shock observations do not appear to change the key results, it is possible that the aid shock effect is only evident in the countries with highest levels of aid dependence. So we now consider the possibility that aid shocks are more likely to lead to violence only in a few influential aid-dependent states. High aid dependence in the developing world is a common occurrence. Svensson (2000) notes that for the fifty most aid-dependent countries the mean value of aid as share of central government expenditures for the period 1975-1995 was 53.8 percent. We find that aid shocks are destabilizing at all levels of aid dependence. We estimate the relationship three different ways, each of which provides more detailed insights into the aid dependence – aid shock relationship.

We measure aid dependence as the logged proportion of aid as a share of GDP—ln(Aid/GDP). As a first cut, we interact this with the variable *Aid Shock* and estimate the rare-events logit as in the primary specifications. We find that this interaction term is statistically insignificant and that it is negative (b = -7.76, p = 0.33), suggesting that, if anything, aid shocks are most destabilizing in states that are not particularly reliant on foreign aid.

A second approach is to split the sample and estimate the effect of aid shocks on conflict onset for states that are in different percentiles of aid dependence: the lower fifty percent (b = 2.28, p = 0.004), the upper fifty percent (b = 0.66, p = 0.015), and the upper twenty-five percent (b = 0.79, p = 0.034). Again we find that the impact of aid shocks on conflict is slightly stronger in less aid dependent countries.

Building on this approach of splitting the sample, a third and more complete way of assessing how the effects of aid shocks change with the level of aid dependence is to estimate models on all subsets of the data defined by the quantiles of aid dependence. We begin with the subset of the data for which the values of aid dependence are between the 10^{th} and 35^{th} percentiles for the whole sample. On this subset, we estimate the effect of aid shocks, plot its coefficient, and include a confidence band defined by the estimated standard error. Then we move the window up, repeating the process for the 11^{th}–36^{th} percentiles of aid dependence, the 12^{th}-37^{th} percentiles, the 13^{th}-38^{th} percentiles, and so on. The key is that as we incrementally increase the level of aid dependence in the subsamples, the effect of aid shocks should increase if higher levels of aid dependence have an amplifying effect on aid shocks. Figure A5 shows the results of these models.

Again, there is no evidence that increased aid dependence increase the effect of aid shocks on conflict onset. We find that the estimated effect of aid shocks is largest for the subsamples of the data that have the least aid dependence, but the confidence bands around these estimates show that there is no statistically significant interaction between *Aid Shocks* and *Aid Dependence*.

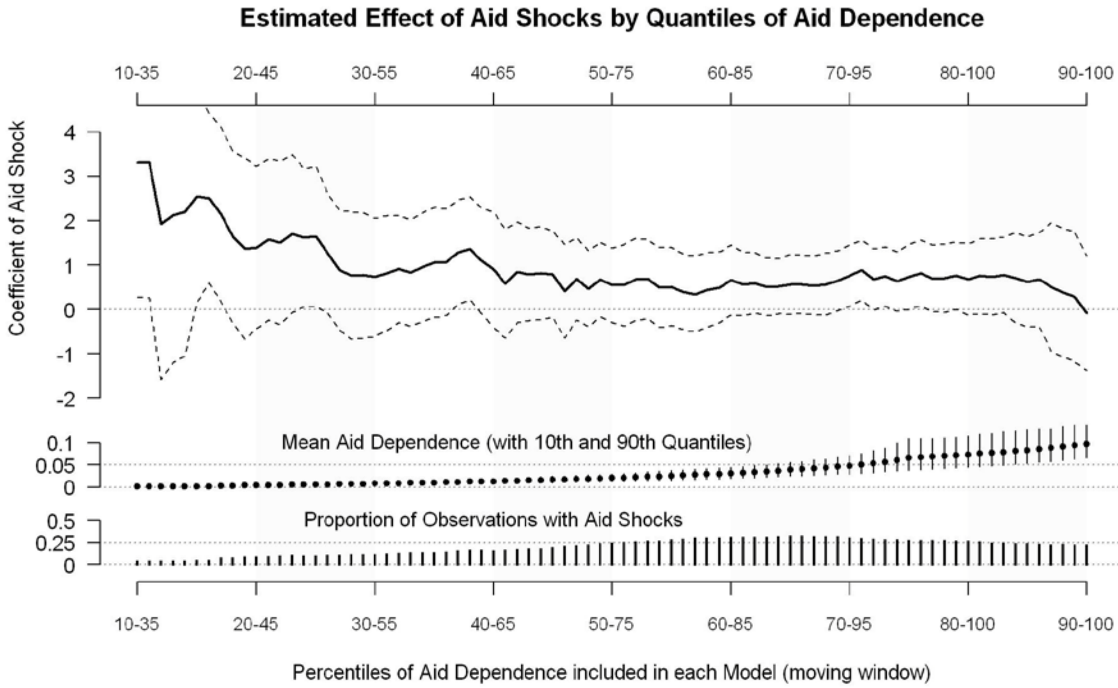

Figure A5: *The estimated coefficient of* Aid Shock *in subsamples of the data defined by percentiles of* Aid Dependence.

A13. Panel Bootstrap

Panel data (such as ours) violate the assumption that observations are i.i.d. (identically and independently distributed). Observations within panels are often dependent in ways that can lead to underestimates of the true variance associated with each coefficient, leading to false statistical significance. One solution to this is the panel bootstrap. We sample 139 panels with replacement from the observed pool of panels and re-estimate the model on each bootstrapped sample. We collect the coefficients for each model and calculate confidence intervals by taking the 2.5th and 97.5th percentiles of the empirical distribution of the coefficients. The results of this procedure for 500 bootstrapped samples are shown in Table A5 and offer robust support for the main findings reported in the paper.

A14. Time, Time^2, and Time^3

Carter and Signorino (2006) suggest the use of the first three polynomials of time as an alternative to modeling time dependence in binary data with cubic splines. The results are qualitatively similar to the main findings and are presented in Table A5.

A15. Omitting 1991-1996 from the sample

Some analysts have noted an apparent increase in civil wars following the end of the Cold War. We find that a relatively large number of aid shocks occurred at the end of the Cold War likely because donors reassessed their client relationships and attempted to recoup a "peace dividend." Although these stylized facts are consistent with our theory, we show that our findings are robust to excluding the years 1991-1996 from the sample in Table A5.

A16. Multiple Imputation

Many observations were excluded from analysis due to list-wise deletion. However, as noted by King et al. (2001), list-wise deletion can introduce bias that can be corrected by multiply imputing the missing observations while preserving uncertainty over their precise values and then estimating models on all multiply imputed datasets simultaneously. We thus multiply impute the missing country-years in our sample using *Amelia II* software (Honaker, King, and Blackwell, 2005-2010) and find similar results. The results are shown in Table A5.

Table A5: Additional Robustness Models

	(1) Panel Bootstrap	(2) Including *time*, *time*^{2}, and *time*^{3}	(3) Excluding 1991-1996	(4) Multiple Imputation
Aid Shock	0.94**	0.85***	1.10***	0.63**
	(0.37, 1.45)	(0.28)	(0.36)	(0.24)
Positive Aid Shock	0.15	0.055	0.27	0.008
	(-0.66, 0.79)	(0.36)	(0.37)	(0.28)
Human Rights Violations	0.62**	0.67***	0.75***	0.48***
	(0.31, 0.90)	(0.14)	(0.18)	(0.13)
Assassinations	0.12	0.17*	0.13	0.16
	(-0.094, 0.34)	(0.10)	(0.13)	(0.089)
Riots	-0.009	0.010	0.10	0.10
	(-0.28, 0.28)	(0.15)	(0.13)	(0.093)
General Strikes	-0.019	-0.022	0.21	0.098
	(-0.51, 0.47)	(0.21)	(0.27)	(0.15)
Anti-Gov. Demonstrations	-0.066	-0.050	-0.20*	-0.082
	(-0.43, 0.10)	(0.13)	(0.11)	(0.075)
Infant Mortality	0.004	0.006	-0.0014	0.005
	(-0.007, 0.014)	(0.005)	(0.0052)	(0.004)
Bad Neighborhood	-0.0476	-0.042	-0.14	0.005
	(-0.33, 0.19)	(0.12)	(0.15)	(0.096)
Partial Autocracy	0.23	0.16	0.22	0.36
	(-0.42, 0.91)	(0.34)	(0.43)	(0.29)
Partial Democracy	-.74	-0.78	-0.66	-0.75**
	(-1.79, 0.17)	(0.50)	(0.60)	(0.35)
Factional Democracy	0.70	0.66	0.58	0.75***
	(-0.014, 1.47)	(0.41)	(0.50)	(0.27)
Full Democracy	0.16	0.055	-0.050	0.25
	(-1.28, 1.16)	(0.59)	(0.69)	(0.35)
ln(GDP per capita)	-0.20	-0.21	-0.32	-0.16
	(-0.74, 0.32)	(0.25)	(0.25)	(0.13)

Continued below.

Table A5 Continued: Additional Robustness Models

	(1) Panel Bootstrap	(2) Including *time*, *time²*, and *time³*	(3) Excluding 1991-1996	(4) Multiple Imputation
ln(Population)	0.10	0.053	0.10	0.052
	(-0.12, 0.34)	(0.089)	(0.10)	(0.057)
Oil	0.009	0.013***	0.010***	0.005***
	(-0.43, 0.74)	(0.28)	(0.37)	(0.20)
Ethnic Frac.	1.40**	1.25**	1.61**	1.04***
	(0.14, 2.95)	(0.61)	(0.76)	(0.37)
Religious Frac.	-0.77	-0.90	-0.21	-0.69
	(-2.22, 0.77)	(0.72)	(0.79)	(0.41)
Non-contiguous	0.98**	1.13***	1.21***	0.88***
	(0.15, 2.03)	(0.34)	(0.34)	(0.24)
Mountains	0.091	0.11	0.12	0.12**
	(-0.11, 0.31)	(0.10)	(0.10)	(0.056)
Cold War	0.18	-1.23**	0.41	0.23
	(-0.56, 0.72)	(0.58)	(0.38)	(0.18)
Spline 1	-0.002		-0.0054	-0.003
	(-0.012,0.004)		(0.0042)	(0.003)
Spline 2	-0.004		0.0071	0.005***
	(-0.003,0.013)		(0.0046)	(0.0007)
Spline 3	-0.002		-0.0042	-0.003***
	(-0.007,0.003)		(0.0029)	(0.0001)
Time		0.20		
		(0.58)		
Time2		-0.0012		
		(0.032)		
Time3		-0.000006		
		(0.0005)		
Constant	-6.15	-7.54*	-5.54*	-5.24***
	(-13.03, 0.89)	(4.31)	(3.22)	(0.60)
Observations	2627	2627	1990	3438

Model 1: Panel bootstrapped 95 percent confidence intervals in standard errors. Based on 500 panel bootstraps.
Models 2-4: Robust standard errors in parentheses. Standard errors clustered by country.
$^{*} p < 0.1$, $^{**} p < 0.05$, $^{***} p < 0.01$.

A17. Multicollinearity

Some of our control variables might be highly collinear with each other, or perhaps with *Aid Shocks*, leading to over-estimates of the standard errors in our regression model. We test for this by calculating variance inflation factors (VIF) for each covariate in our primary specification. The standard advice is that when variance inflation factors exceed 5, the variable might be too collinear with other variables. From Table A6 below, we can see that none of the variance inflation factors exceed five, but that *infant mortality* and *ln(GDP per capita)* come close with variance inflation factors above four. Our findings about the effect of aid shocks are entirely robust to excluding *infant mortality* ($b = 0.90$, $p = 0.001$), *ln(GDP per capita)* ($b = 0.99$, $p < 0.001$), or both ($b = 1.08$, $p < 0.001$).

Table A6: Variance Inflation Factors

Variable	VIF	Variable	VIF
Aid Shock	1.29	Factional Democracy	1.85
Positive Aid Shock	1.29	Full Democracy	3.43
Human Rights Violations	2.06	ln(GDP per capita)	4.83
Assassinations	1.12	ln(Population)	1.80
Riots	1.76	Oil	1.22
General Strikes	1.24	Instability	1.15
Anti-Gov. Demonstrations	1.82	Ethnic Frac.	1.88
Infant Mortality	4.09	Religious Frac.	1.31
Bad Neighborhood	1.49	Non-contiguous	1.46
Partial Autocracy	1.83	Mountains	1.22
Partial Democracy	2.11	Cold War	1.26

Variance inflation factors (VIF) for covariates included in the primary specification. Inflation factors greater than five are often considered problematic.

Variable Appendix

A18. Variable Definitions and Sources

Ethnic fractionalization: Drawn from Fearon and Laitin (2003), this variable measures ethnolinguistic fractionalization, meaning "the probability that two randomly drawn individuals in a country are from different ethnolinguistic groups" (78).[4]

Religious fractionalization: Similar to ethnic fractionalization, this variable (Fearon and Laitin 2003) measures the probability of two randomly drawn individuals in a country being from different religious groups.

Oil: We use the measure of oil from Ross (2008), which captures oil rents per capita. A full explanation of the variable can be found in the Appendix of Ross (2008).

Instability: This variable is taken from Fearon and Laitin's (2003) data. They created a dichotomous variable representing whether a country had a change of three or greater on the Polity IV scale in the last three years (81).

Population: This variable is the natural log of the total population of each country. We drew this variable from Gleditsch's Expanded Trade and GDP Data (2004).

Noncontiguous state: We drew this variable from Fearon and Laitin (2003), who define it as "countries with territory holding at least 10,000 people and separated from the land area containing the capital city either by land or by 100 km of water" (81).

Mountains: This variable measures the percent of a country that is mountainous. We employ Fearon and Laitin's (2003) variable in which they used A.J. Gerard's data as their base and then filled

[4] For more information on the variables, please see the cited references.

in missing values using CIA World Factbook figures. They then took the log of the values to create their final variable.

Democracy: We use the five-category measure of democracy introduced (and described more fully) by Goldstone et al (2010) because they find that it is important for predicting political instability and state failure. This measure divides regimes into five types based on the Polity IV scales of *Executive Recuritment* (EXREC) and *Competitiveness of Political Participation* (PARCOMP). Below, we reproduce Figure 1 of Goldstone et al to show the exact construction of these variables.

- **Full Autocracy:** These regimes "combine an absence of effective contestation for chief executive with repressed or suppressed political participation" (195).
- **Partial Autocracy:** These regimes either allow elections for national office but repress participation *or* allow some participation but restrict candidates for office.
- **Partial Democracy:** These regimes have both competitive executive recruitment and substantial participation, "but either elections are not fully free and fair, or political participation is not fully open and well institutionalized" (196).
- **Factional Democracy:** This is a specific sub-type of partial democracy that exhibits factionalism – "sharply polarized and uncompromising competition between blocs pursuing parochial interests at the national level" (196).
- **Full Democracy:** Characterized by free and fair elections and open and robust political participation.

	Competitiveness of Political Participation					
Executive Recruitment	Repressed (0)	Suppressed (1)	Unregulated (2)	Factional (3)	Transitional (4)	Competitive (5)
(1) Ascription						
(2) Ascription + Designation						
(3) Designation						
(4) Self-Selection						
(5) Transition from Self-Selection						
(6) Ascription + Election						
(7) Transitional or Restricted Election						
(8) Competitive Election						

White: Full Autocracy; Light grey: Partial Autocracy; Dark grey: Partial Democracy; Very dark grey: Factional Democracy; Black: Full Democracy.

In alternative specifications, we also used the more common 21-point democracy scale created using the Polity IV data (Marshall et al. 2008). We employed Polity's measure in which they convert their standardized authority scores (i.e. -66, -77, -88) into standard Polity scores and then lagged the variable by one year.

GDP per capita: We employed a one-year lag of Gleditsch's (2004) GDP per capita data based on his version 4.1. In that version, he uses Penn World Tables 6.1 data as his base data and then fills in missing values either using the CIA World Factbook figures or else by interpolation. Gleditsch calculated his variable in real figures using constant U.S. dollars (base 1996). For a more thorough explanation of his data, please see Gleditsch (2002).

Cold War: We measure this variable dichotomously, coding all years during the Cold War as 0 and all years after the Cold War as 1. We determined that the Cold War ended in 1991 and code that year as a 0.

Human Rights Violations: We use the Political Terror Scale (Gibney, Cornett, and Wood 2010), which codes human rights violations on a 1 to 5 scale, where 1 indicates very rare or non-existent human rights violations and 5 represents severe, population-wide human rights violations, including torture, political killings, disappearances, etc.

Assassinations, Riots, General Strikes, and **Anti-government Demonstrations:** These four variables are all measured as counts of events in the *Cross-National Time-Series Data Archive* collected by Banks (2008). Assassinations are defined as "any politically motivated murder or attempted murder of a high government official or politician" (Users Manual, 11). Riots are defined as "any violent demonstration or clash of more than 100 citizens involving the use of physical force" (12). General Strikes are defined as "any strike of 1,000 or more industrial or service workers that involves more than one employer and that is aimed at national government policies or authority" (11). Anti-government Demonstrations are "any peaceful public gathering of at least 100 people for the primary purpose of displaying or voicing their opposition to government policies or authority, excluding demonstrations of a distinctly anti-foreign nature" (11).

Infant Mortality: This variable is taken from the World Bank's *World Development Indicators* (2008) and measures the number of deaths among children less than a year old (per 1000 live births).

Bad Neighborhood: This measures the number of neighboring countries that are experiencing a civil or ethnic war (Marshall 2010).

Data References

Banks, Arthur S. 2008. "Cross-National Time-Series Data Archive." Accessed July 7, 2009 via Harvard Library.

Carter, David, and Curt Signorino. 2006. "Back to the Future: Modeling Time Dependence in Binary Data." Unpublished Manuscript. University of Rochester.

De Ree, Joppe, Eleonara 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: 301-313.

Deaton, A.. 2009. "Instruments of Development: Randomization in the Tropics, and the Search for the Elusive Keys to Economic Development," NBER Working Paper #14690.

Easterly, William. 2009. "Empirics of Strategic Interdependence: The Case of the Racial Tipping Point," *The B.E. Journal of Macroeconomics*: Vol. 9 : Iss. 1. Article 25.

Fearon, James D., and David Laitin. 2003. "Ethnicity, Insurgency, and Civil War." *American Political Science Review* 97 (1): 75-90.

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

Gibney, Mark., Linda Cornett, and Reed Wood. 2010. *Political Terror Scale 1976-2006.* Accessed March 2010 at <http://www.politicalterrorscale.org/>.

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

Gleditsch, Kristian Skrede. 2004. "Expanded Trade and GDP Data." Available at <http://privatewww.essex. ac.uk/~ksg/exptradegdp.html> (July 22, 2008).

Goldstone, Jack A., Robert H. Bates, David L. Epstein, Ted Robert Gurr, Michale B. Lustik, Monty G. Marshall, Jay Ulfelder, and Mark Woodward. 2010. "A Global Model for Forecasting Political Instability." *American Journal of Political Science* 54 (1): 190-208.

Heckman, James. 1995. "Instrumental Variables: A Cautionary Tale." NBER Technical Working Paper No. 185. Available at <http://www.nber.org/papers/t0185.pdf>.

Honaker, James, Gary King and Matthew Blackwell. 2005-2010 "Amelia II: A Program for Missing Data" . <http://gking.harvard.edu/amelia/>.

King, Gary, James Honaker, Anne Joseph, and Kenneth Scheve. 2001. "Analyzing Incomplete Political Science Data: An Alternative Algorithm for Multiple Imputation." *American Political Science Review* 95 (1): 49-69.

Kuziemko, Ilyana and Eric 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): 905-93.

Marshall, Monty G. 2010. "Major Episodes of Political Violence (MEPV) and Conflict Regions, 1946-2006." Center for Systemic Peace. Accessed February 17, 2010 at <www.systemicpeace.org>.

Marshall, Monty G., Keith Jaggers, and Ted Robert Gurr. 2008. "Polity IV Project: Political Regime Characteristics and Transitions, 1800-2006." Available at <http://www.systemicpeace. org/inscr/inscr.htm> (July 21, 2008).

Miguel, Edward, Shankar Satyanath and Ernest Sergenti. 2004. "Economic Shocks and Civil Conflict: An Instrumental Variables Approach." *Journal of Political Economy* 112 (4): 725-753.

Murray, Michael P. 2006. "Avoiding Invalid Instruments and Coping with Weak Instruments." *Journal of Economic Perspectives* 20 (4): 111-132.

Neumayer, Eric 2003c. *The Pattern of Aid Giving: The Impact of Good Governance on Development Assistance.* London: Routledge.

Nielson, Daniel L. and Michael J. Tierney. 2005. "Theory, Data, and Hypothesis Testing: World Bank Environmental Reform Redux." *International Organization* 59 (3): 785-800.

OECD. 2010. “User’s Guide to the CRS Aid Activities Database.” Accessed at < http://www.oecd.org/document/50/0,3343,en_2649_34447_14987506_1_1_1_1,00.html >, 5 May 2010.

Powell, Robert. 2004. “The Inefficient Use of Power: Costly Conflict with Complete Information.” *American Political Science Review* 98 (2): 231-241.

Rosenbaum, Paul R. and Donald B. Rubin. 1983. “The Central Role of the Propensity Score in Observational Studies for Causal Effects.” *Biometrika* 70 (1): 41-55.

Ross, Michael. 2008. “Oil, Islam, and Women,” *American Political Science Review*, 102 (1): 107-23.

Sambanis, Nicholas. 2004. “What Is Civil War? Conceptual and Empirical Complexities of an Operational Definition.” *Journal of Conflict Resolution* 48 (6): 814–858.

Svensson, J. 2000. "Foreign Aid and Rent-Seeking," *Journal of International Economics*, 51(2), 437-461.

Stock, James H., Jonathan H. Wright, and Motohiro Yogo. 2002. “A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments.” *Journal of the American Statistical Association* 20 (4): 518-529.

USAID. 2010. *U.S. Overseas Loans and Grants: Obligations and Loan Authorizations, July 1, 1945-September 20, 2008*. Accessed April 22, 2010 at < http://gbk.eads.usaidallnet.gov/data/ >

World Bank. 2008. “World Development Indicators.” Available at <http://web.worldbank.org/data>.