ISSN 2042-2695

CENTRE *for* ECONOMIC PERFORMANCE

CEP Discussion Paper No 1694

May 2020

Damned by Dams? Infrastructure and Conflict

Ulrich J. Eberle

THE LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE

Economic and Social Research Council

Abstract
This study investigates the impact of dams on local conflict across the world with georeferenced location information on dams and conflict events for the years 1989 to 2016. The identification strategy exploits exogenous variation in the river gradient to instrument for endogenous dam placement. The results document strong and robust evidence for an increase in intrastate conflict in the immediate vicinity of newly-built dams, but no robust effect for interstate conflict can be identified. Examining the mechanisms, ethnically polarized and fractionalized regions are more likely to experience the negative economic consequences as well as a surge in violence associated with dams. Further, countries with low levels of political competition are subject to more violence, suggesting that an institutional failure to account for local preferences may lead to violent confrontations. Finally, the policy analysis reveals that organizations providing the funding for dams, usually international financial institutions, have effective tools such as transboundary water treaties to prevent an outbreak of violence by enforcing regulation and monitoring during the implementation phases of dams.

Key words: Dams, Infrastructure, Conflict, Civil War, Ethnicity, Fractionalization, Inequality
JEL Codes: D74; H54; N40; O13; O18; Z18

This paper was produced as part of the Centre's Urban and Spatial Programme. The Centre for Economic Performance is financed by the Economic and Social Research Council.

I am grateful to Steve Gibbons, Vernon Henderson, Rafael Lalive, Guy Michaels, Dzhamilya Nigmatulina, Henry Overman, Paolo Pinotti, Dominic Rohner and Mathias Theonig and thank participants at various conferences and seminars for their valuable comments. I acknowledge financial support from the Swiss National Science Foundation Doc.Mobility fellowship P1LAP1_181253 and especially thank Vernon Henderson for his hospitality at the London School of Economics during 2018 to 2020, when the main work for this paper was completedacknowledges support from the French National Research Agency Grant (ANR-17-EURE-0020).

Ulrich J. Eberle, University of Lausanne and Centre for Economic Performance, London School of Economics.

Published by
Centre for Economic Performance
London School of Economics and Political Science
Houghton Street
London WC2A 2AE

1 Introduction

Economic development is strongly associated with infrastructure investment. Wealth generated from infrastructure projects is however not necessarily evenly spread across space and tends to disadvantage communities in their immediate vicinity. In many instances, directly affected communities oppose planned projects, as they fear displacement, economic hardship, environmental degradation and other negative externalities caused by a new dam, highway or airport. Dams are a particularly frequent source of disputes, with numerous cases of peaceful protests escalating into violent confrontations with government forces. Despite controversies, the world economy heavily relies on dams. By 2015, dams produced 16 percent of the global electricity and 35 countries derived at least half of their energy supply from this renewable source (World Bank, 2019). Dams further accounted for an estimated 12 to 16 percent of the global food production, with one third of the global irrigated agriculture relying on water from reservoirs in the year 2000 (World Commission on Dams, 2000). The dependency on dams is expected to increase in the future, as the world faces the beginning of a global dam building boom, with two trillion USD investments in hydropower projects alone over the next decade (Zarfl et al., 2015).

Although their overall importance is undeniable, dams can put severe pressure on directly affected communities through physical displacement and income losses. The ongoing trend towards ever larger projects is a reason for particular concern (c.f. Figure A1 in the Appendix). For example the Chinese Three Gorges Dam completed in 2016, the largest hydropower project in the world, alone displaced at least 1.2 million people (Exchange on Environment Conflict and Cooperation, 2019). A growing body of qualitative research accuses dams of fueling intrastate conflict and soaring interstate tensions, with displacement and inadequate compensation of economic losses at the core of the problem. For instance, during the construction of the Chixoy dam in Guatemala in the early 1980s, 444 Maya Achi minorities were massacred by army and paramilitary forces, because they refused to evacuate their homes located in the planned catchment area of the dam. The villagers, mainly women and children, feared to lose their livelihoods, inadequate compensation by the government and resettlement to arid highlands (ICTA-UAB, 2019). While the Chixoy dam exhibits an extreme degree of brutality it is by far not an isolated incident of dam-related violence, but rather exemplifies numerous anecdotal evidence along similar lines. To determine whether a general pattern of dam-induced violence can be observed, quantitative evidence is necessary to assess the underlying mechanisms and preventive policies.

This paper is, to the best of my knowledge, the first global empirical investigation into the impact of dams on conflict, aiming at providing causal evidence and a profound understanding of the role of political and ethnic power relations. Dams are an especially suitable type of infrastructure for empirical analyses, because unlike most of the benefits, the negative externalities of dams are highly localized and mostly bound to their immediate vicinity. The empirical strategy is based on a panel dataset covering most parts of the world for the years 1989 to 2016. Dam data is from the Global Reservoir and Dam database, a high-quality data monitoring the exact location, type, size and completion date of large dams worldwide and conflict data comes from the Uppsala Conflict Data Program - Georeferenced Event Dataset, depicting the exact location, year and type of conflict incidents (Lehner and Grill, 2013; Sundberg and Melander, 2013). The impact of dams is analyzed at the subnational level, with each country further split into *drainage basins*, a spatial unit based on elevation data which allows to define watershed boundaries and evaluate the impact of dams.[1]

One major difficulty in identifying the causal effect of dams on conflict is that governments may not allocate dams randomly within their countries, but rather prefer to locate them in peaceful and economically thriving regions. To overcome this potential source of endogeneity, I employ an instrumental variable strategy in the fashion of Duflo and Pande (2007), which exploits the fact that it is cost efficient - from an engineering point of view - to place dams in areas with a favorable river gradient. Natural concerns about the exclusion restriction are addressed with the empirical specification accounting for a set of controls, drainage basin fixed effects and country-year fixed effects. In addition, several robustness checks and placebo tests are performed to further assess the validity of the instrument.

The first part of the empirical analysis investigates the impact of dams on local conflict in the instrumental variable setting. The results show an substantial increase in conflict, especially in intrastate conflict incidents involving up to 50 casualties. The results remain robust

[1] Please consult Section 4 for an extended discussion of drainage basins.

when controlling for potentially confounding geographic factors, the distance to national borders, as well as when accounting for the lagged dependent variable and basin-specific time trends. Quantifying the baseline effect shows that currently planned dams are projected to increase the frequency of global conflict incidents by up to 5.3 percent by 2030. To study the timing of conflict onset, I perform a synthetic control method. Results confirm the baseline findings and further show that conflict may already start during the final years of dam construction, with an escalation of violence after the opening of dams. Analyzing the involved fighting parties, I find no statistically significant association between dams and interstate conflict. This does not preclude an increase in geopolitical tensions due to dams, but rather rules out an escalation of transboundary disputes. Consistent with the domestic nature of violence, the analysis further documents that conflict does not appear to diffuse to other upstream or downstream regions. Heterogeneity can be identified with respect to certain country characteristics: violence stems from countries with large agricultural sectors, whereas neither the average per capita income nor the degree of democratization in a country appears to matter.

The second part of the analysis studies the channels of transmission of violence. The economic impact of dams is assessed with disaggregated GDP data and satellite night-time lights data (Nordhaus, 2006; National Oceanic and Atmospheric Administration, 2013). Results show some evidence for a sizable, negative effect to the local economy in regions with new dams. A large shock to the local economy appears plausible, because affected economic centers, usually located along rivers banks, are now replaced by reservoirs. This finding may suggest that income losses reduce the opportunity costs for individuals to join rebel forces, which would provide a plausible explanation for the surge in violence. Next, I test institutional characteristics likely to result in the mismanagement of dams and find that conflict is more pronounced in countries with low levels of political competition. Those countries are less likely to represent the preferences of affected communities, which I argue could lead to an escalation of violence. Further, the role of ethnic grievances is addressed at the basin-level. The estimates document that new dams cause larger economic damages and at the same time amplify conflict frequency in highly polarized and fractionalized regions.

The final part of the analysis considers institutional features and policy interventions likely to mitigate conflict. I first test whether strong judicial systems may help to enforce the rights of affected communities and results show some evidence for a reduction in conflict in countries with independent judiciaries. Moreover, international financial institutions which usually provide the funding of dams, have regulatory tools such as transboundary water treaties that can effectively prevent an outbreak of violence.

This paper contributes to the literature in several ways. It is, to the best of my knowledge, the first paper empirically studying the link between dams and conflict - or even more broadly between infrastructure and conflict - at a highly disaggregated level, considering all types of large dams. Covering most parts of the world ensures the external validity of the results and also allows the analysis of cross-national heterogeneity in the sensitivity to conflict. The findings are in line with the qualitative literature and further provide novel insights that open new avenues to understand the implications of infrastructure spending and to prevent conflict in the future.

The remainder of the paper is organized as follows: The next section surveys the related literature. Section 3 provides background information and Section 4 presents the data and descriptive statistics. Section 5 introduces the identification strategy and presents the main results. Sections 6 and 7 report further sensitivity checks and heterogeneous effects, before moving on to the analysis of the mechanisms in Section 8. Section 9 addresses policies mitigating conflict and Section 10 concludes.

2 Literature Review and Conceptual Framework

This paper contributes to two main strands of literature. The first strand evaluates the local economic impact of dams and other infrastructures. Duflo and Pande (2007) study the distributional effects of large irrigation dams in India. This landmark paper pioneers an instrumental variable approach exploiting the predictive power of the river gradient to instrument for dams placement. They find an increase in agricultural productivity and a decrease in poverty in districts downstream of dams. In districts containing dams, poverty and productivity increase, whereby the latter being statistically insignificant. Strobl and Strobl (2011) repeat the exercise at the drainage basin level for African irrigation dams and employ a similar identification strategy. They confirm the increase in agricultural productivity downstream of dams, but find a decrease in productivity in the vicinity of dams. Their study is limited the regions subject to transboundary water treaties, which I show in Section 9 are less likely to experience conflict.

More distantly related is the literature evaluating other types of infrastructure, such as highways (Chandra and Thompson, 2000; Michaels, 2008) and railroads (Atack et al., 2010; Banerjee, Duflo and Qian, 2012). Further, indirectly related is the literature analyzing the effect of electrification on a number of outcome variables. For instance, Dinkelman (2011) demonstrates how electrification increases female employment in South Africa. Lipscomb, Mobarak and Barham (2013) find an overall positive effect of electrification on economic development in Brazil. On the local level, agricultural productivity initially drops before the effect reverses in counties where hydropower dams have been placed.

The second strand of literature studies conflict. The empirical strategy of this paper is closely related to the recent series of work studying the causes of conflict at a highly-disaggregated level (E.g. Michalopoulos and Papaioannou, 2013; Berman et al., 2017; Harari and La Ferrara, 2018; Eberle, Rohner and Thoenig, 2019). These papers frequently employ panel data at the grid cell level, in order to analyze the dynamics of conflict at the subnational level. The disaggregated nature of these studies is appealing from an identification point-of-view, as it allows for more complex fixed effect structures capturing not only omitted variables across countries, but also time-invariant heterogeneity within countries.

This paper is conceptually closely related to a large body of work investigating drivers of conflict. Dams and their associated reservoirs eliminate parts of the local economy, by reducing the supply of crop land and displacing affected communities. Conflict theory predicts two possible outcomes: income losses of affected communities could lower the *opportunity costs* of rebellion. In other words, the likelihood of conflict may increase when individuals have now "less to lose" when exchanging farm labor for joining rebel forces. Conversely, if dams cause incomes from agriculture to fall, the potential gains from *appropriating* farm land are reduced, which in turn could lead to a reduction in violence (Miguel, Satyanath and Sergenti, 2004; Bazzi and Blattman, 2014). The above logic can be reversed in case crop land becomes more valuable, e.g. through better access to water from irrigation facilities. In addition, dams frequently degrade soil and endanger biodiversity, which may fuel anti-state grievances (Berman et al., 2017).

One could argue that violence is not caused by dams per se, but rather by fragile institutions. *Weak states* are notoriously associated with conflict (Fearon, 2005; Besley and Persson, 2011). Notably, extractive institutions could lead to an unequal distribution of wealth generated by dams. Duflo and Pande (2007) argue that Indian states with extractive institutions may be responsible for an increase in poverty after dam completion, because those states fail to redistribute the gains from dams. Further, Besley, Persson and Sturm (2010) show theoretically and empirically, for the US, that a lack in political competition can lead to growth-suppressing policy outcomes. In the light of this paper, constraints to the democratic decision-making process could result in dams failing to represent the preferences of affected communities. As a result, inadequate compensation for economic losses could directly reduce a government's ability to mitigate conflict (Berman, Shapiro and Felter, 2011). Finally, this paper contributes to the literature studying the impact of ethnic grievances on economic development, public good provision and conflict (Montalvo and Reynal-Querol, 2005; Alesina and Zhuravskaya, 2011; Esteban, Mayoral and Ray, 2012; Alesina, Michalopoulos and Papaioannou, 2016).

The existing literature has focused on evaluating the economic impact of dams, but neglected to consider conflict as a key outcome variable, despite numerous qualitative evidence. Further, this paper is the first work evaluating the local impact of dams worldwide, studying the precise timing of conflict onset. Finally, the role of institutional characteristics and ethnic grievances as a cause of mismanaged infrastructure projects have been widely overlooked and are discussed in detail in this paper. The next section discusses the existing case study literature on the link between dams and violence.

3 Background: Dam Construction in Past Decades

For decades, dams have been built and funded without much consideration for effectiveness and impact, despite their high costs. With the start of the dam building boom in the 1950s, an opposition to dams has begun to form, arguing that the mostly macroeconomic benefits are overshadowed by disruptive and lasting interventions into local ecological and human systems (World Commission on Dams, 2000). By 2015, 43 percent of the global river flow had already been regulated by man-made structures. Currently constructed and planned dams are likely to double river fragmentation by 2030 (Grill et al., 2015). This is problematic because dams block sediments, nutrients and the migration of aquatic organisms, which can severely reduce fish stocks and the soil quality of the surrounding farm land (World Commission on Dams, 2000). Regulated water flows further shift the availability of water, reduce the local water supply and disrupt traditional flood recession agriculture which relies on seasonal floodplains and wetlands. Beyond the supply of water, the supply of agricultural land is also affected because a considerable amount is lost to the command area, irrigation facilities and the reservoirs, frequently inundating the economic and cultural centers located along the river banks.[2] The average reservoir alone covers 65 km^2 of land in my data. As a result, stress put on the environment and society can translate into the displacement of people and their livelihoods.

According to the World Commission on Dams, at least 40 million people have been physically displaced as a consequence of dams by the year 2000. Displacement is observed in more than half of the studied cases, regularly accompanied by a lack or a delay in adequate compensation.[3] And in cases where relief schemes are installed, they tend to favor directly, physically displaced people and land owners with resettlement schemes and cash, but often neglect to compensate indirect economic losses and minorities. In several cases, indigenous, landless and undocumented people did not qualify for compensation schemes, as marginalized groups tend to receive less public recognition and support for their interests, despite being disproportionally exposed to dams. For instance in India, ethnic minorities account for 40 to 50 percent of all physical displacement by dams, while only representing eight percent

[2] The Gilgel Gibe III Dam (Ethiopia): disputes over the impact on traditional flood recession agriculture caused local dispute; The Kalabagh Dam (Pakistan): one of many examples where shifting water resources caused conflict in neighboring regions; Hasankeyf (Turkey) and Lake Turkana (Ethiopia): examples where dams threaten to destroy world heritage sites (Exchange on Environment Conflict and Cooperation, 2019; The Economist, 2019).

[3] Bargi dam (India): 90 percent of the displaced were not resettled and 162 villages were inundated without warning; Yacyreta project (Argentina, Paraguay): 70 percent of the displaced were not resettled 2 years prior to completion; Tarbela dam (Pakistan): 30 percent of the local population, or 20,000 people, remained without compensation (World Commission on Dams, 2000).

of the total population (World Commission on Dams, 2000). All those existential threats to physical, mental and economic well-being may translate into despair, grievances and violence. A majority of violence associated with dams can be classified as intrastate conflict, with government forces fighting against rebels, protesters and civilians.[4] Interstate tensions on the other hand may also rise, when dams planned on transboundary rivers threaten to reduce downstream water supplies.[5]

Dams are notoriously associated with environmental and social issues, it is however important to recognize that dams could contribute to a reduction in conflict. Irrigation and hydropower dams have positive, macroeconomic effects, with hydropower contributing 75 percent to the global renewable energy supply by 2015 (World Bank, 2019). Further, flood and drought control dams are constructed with the purpose to stabilize regions exposed to erratic weather patterns. While their benefits are well-known, the most central issue remains: dams unevenly distribute benefits and negative externalities, resulting in spatial inequality, with populations in the vicinity of dams bearing the majority of the burden.

Timing: Dam construction and local violence

This section consults qualitative evidence from around the world to identify generalizable patterns relating to the onset of violence along the life cycle of dams.[6]

Dam construction: With the start of the construction of a dam, the magnitude and full environmental and social impact becomes apparent. People learn to which extent they will be displaced and local leaders negotiate with the central government over compensating economic losses of affected communities, if not already done. In some cases, grievances exac-

[4] There are numerous examples of disputes and conflict over dams around the world. Recent incidents include the Santa Rita Dam in Guatemala, the Barro Blanco Dam in Panama and several projects in Myanmar. Myanmar exhibits a long list of mismanaged dams and related violence. One example is the recently commissioned Upper Paunlaung Dam. Another example is the when in 2011, violent protests broke out over the Hatgyi Dam and further government's plans to construct six hydro-electric dams along the Salween River in 2013 (Exchange on Environment Conflict and Cooperation, 2019; ICTA-UAB, 2019).

[5] Examples include increasing tensions among multiple nations sharing access to the Mekong river; the Grand Ethiopian Renaissance Dam, which triggered an diplomatic crisis with Egypt located further downstream on the Nile; water disputes between Afghanistan and Iran along the Helmand River; and between Kyrgyzstan and Uzbekistan over the construction of the Kambar-Ata-1 Dam on the Naryn River in central Kyrgyzstan (The Economist, 2016; Exchange on Environment Conflict and Cooperation, 2019).

[6] Case study evidence mainly stems from three sources: World Commission on Dams (2000), Exchange on Environment Conflict and Cooperation (2019) and ICTA-UAB (2019). Section 6.2 empirically assesses the onset of violence.

erbate and peaceful protests already turn violent, as reservoirs may already start to be filled, affecting the local supply of water and land.[7] First relief schemes, if installed, compensate and resettle affected groups. The economic impact still remains somewhat limited, because the dam construction site frequently absorbs a large number of unskilled local workers.[8]

Dam completion: The full economic impact unfolds at this stage. Affected groups located in the catchment area of dams lose not only their land and homes, but also their day laboring jobs at the construction site. Further compensation schemes, if installed, are designed to absorb some of the economic shock. Further, shifting local water supplies start to permanently transform communities beyond the dams catchment area.

4 Data and Descriptive Statistics

4.1 Spatial unit of observation: drainage basins

Given the apparent local nature of conflict associated with dams and to be able to exploit within country variation, it is necessary to subdivide countries further into smaller geographic units. One major drawback of using subnational administrative boundaries such as state borders is that they frequently coincide with rivers. Consequently, a correct assignment of dams to a single state may be ambiguous (Strobl and Strobl, 2011). Further, subnational boundaries should be treated with caution, because state fragmentation might be a direct result of, and hence potentially is endogeneous to, conflict.[9]

To overcome the aforementioned issues, I subdivide countries into *drainage basins*, a globally consistent geospatial unit frequently employed in environmental and hydrology studies. A basin can be best imagined as the land area that collects and drains off precipitation,

[7]For instance in 1978, police forces shot 4 protesters opposing the Chandil dam in India; in 1980, at least 23 protesters opposing the Bakolori dam were killed (World Commission on Dams, 2000, p.18). More recently, the Upper Paunglaung Dam hydroelectricity project completed in 2015 in Myanmar faced strong local resistance from farmers and indigenous groups during the implementation phase over issues of misinformation, mismanagement and inadequate compensation. The 61 km^2 large reservoir submerged 23 villages and displaced 10,000 people (ICTA-UAB, 2019).

[8]For instance during the construction of the Kariba dam on the Zambezi river, 10,000 local unskilled laborer were hired (World Commission on Dams, 2000).

[9]Recent contributions studying the effects of subnational boundary changes include Lipscomb and Mobarak (2011); Swee (2015); Grossman, Pierskalla and Boswell Dean (2017).

such as a valley. Drainage basin data is from the HydroBASINS dataset by the Conservation Science Program of the World Wildlife Fund and USGS (Lehner and Grill, 2013). The data is part of the HydroSHEDS environment and can seamlessly be combined with further useful information on elevation and the river network, which is used throughout this paper. Drainage basins are generated with an elevation-based algorithm that breaks out sub-watersheds from larger river basins. One useful by-product is that rivers are usually located towards the center of drainage basins, enabling an unambiguous assignment of dams to a single basin. Panel A of Figure A2 in the Appendix shows an example of a basin containing the Upper Paunlaung gravity dam on the Paunlaung river in Myanmar. My sample consists of 15,923 drainage basins, covering the world below 60 degrees latitude.[10] The HydroBASIN algorithm generates drainage basins that are comparable in size and the average basin measures an area of 6,903 km^2, which appears to be a usable size to analyze local conflict. Countries with a single drainage basin are excluded from the sample, since those observations would be absorbed by the country-year fixed effects. The average (median) country consists of 104 (35) drainage basins and in cases where basins range across national borders, the basins are split along the border to avoid an incorrect assignment of dams or conflict to the wrong country. Finally, it is important to mention that each drainage basin records the upstream and downstream neighboring basins, allowing to empirically address whether dams diffuse violence.

4.2 Conflict data

Conflict outcome variables are drawn from the Uppsala Conflict Data Program - Georeferenced Event Dataset (UCDP - GED), a geocoded conflict dataset for the years 1989 to 2016 (Sundberg and Melander, 2013). The data is particularly suitable for the study at hand, because it covers almost the entire world and is available for a relatively long period.[11] Con-

[10]Data above 60 degrees latitude is not available from a single data source in HydroSHEDS, because of increasing distortions towards the poles. It turns out the omitted areas contain only 12 dams built since 1989 (most of them located in Norway and Russia) and less than .5 percent of the global conflict, hence unlikely to impair the external validity of this study. Drainage basin data is available on 12 aggregation levels. I follow Strobl and Strobl (2011) and use level 6, usually corresponding to subsections of larger river basins. For instance, the more than 4,000 km along the Mekong river are divided into 65 basins.

[11]Syria is the only country not covered by UCDP GED, because no reliable data could be collected. Alternative georeferenced datasets include Armed Conflict Location and Event Data Project (ACLED) and

flict event information is collected from multiple sources such as global news, local media monitored by the BBC and secondary sources such as NGOs and field reports. The data collection process is subject to multiple quality checks and validation protocols to minimize misreporting. Coverage may still vary across countries and time, due to differences in censorship, the freedom of press or media interest in specific conflicts, underlining the necessity of country-year fixed effects in the empirical strategy. The data registers conflict events with a minimum of one fatality. In addition, only conflict events of armed groups are considered if the group is associated to a minimum of 25 deaths in at least one of the sample years. For each event, the exact location, along further useful variables such as the involved fighting parties, the number of deaths and the type of conflict, is documented. Events are geocoded with high precision and location information is accurate to the village level, hence sufficient to assign each event to a drainage basin.

Measurement error may occur when events remain unreported or imprecisely recorded with respect to their location (König et al., 2017). Unlike ACLED, UCDP - GED focuses on established perpetrators, which minimizes the likelihood of missing or poorly documented events. Further, different types of conflict may be recorded with different precision (e.g. media coverage of large-scale battles tends to be better documented than smaller, localized disputes). The sensitivity analysis section exploits information on the type of conflict and the involved fighting groups and finds that the results are consistent throughout. In most specifications, I aggregate conflict events to the basin-year level and define a binary outcome variable taking the value one if at least one event occurs in a basin in a given year, zero otherwise. Conflict dummies are a conventional way to minimize a potential non-classical measurement error stemming from over-reporting of larger battles.

4.3 Dam data

Data on dams is taken from the third version of the Global Reservoir and Dam Database (GRanD) (Lehner et al., 2011). GRanD combines, harmonizes and cross-validates multiple existing national and continental dam registers into a single, globally consistent product.

Global Data on Events, Language, and Tone (GDELT). ACLED is available for Africa from 1997 onward and other regions of the world are available for recent years only.

Unlike alternative global dam data, GRanD is the only data providing the precise geocoding of dams and their reservoirs in addition to a detailed list of attributes, making it the optimal candidate for this study.[12] The list of attributes include specifications for each dam (height, length, main and secondary use) and the according reservoir (capacity, surface, upstream catchment capacity). Most importantly, the year of dam completion is recorded, which allows to observe conflict before and after the presence of dams in a basin.[13] The completion year is reported for 95 percent of all dams, and I additionally collected data on the construction start year. The average (median) dam in my sample takes 7.5 (5) years to be built. It is important to note that, due to limitation of some of the underlying national registers, GRanD only records relatively large dams with a reservoir size of at least 0.1 km^3. This may limit the external validity of the analysis, but at the same time avoids putting too much weight on very small dams. Finally, manual inspection shows a high degree spatial precision, as depicted for the aforementioned Paunlaung Dam in Myanmar in Figure A2, Panel B in the Appendix.

4.4 Descriptive statistics

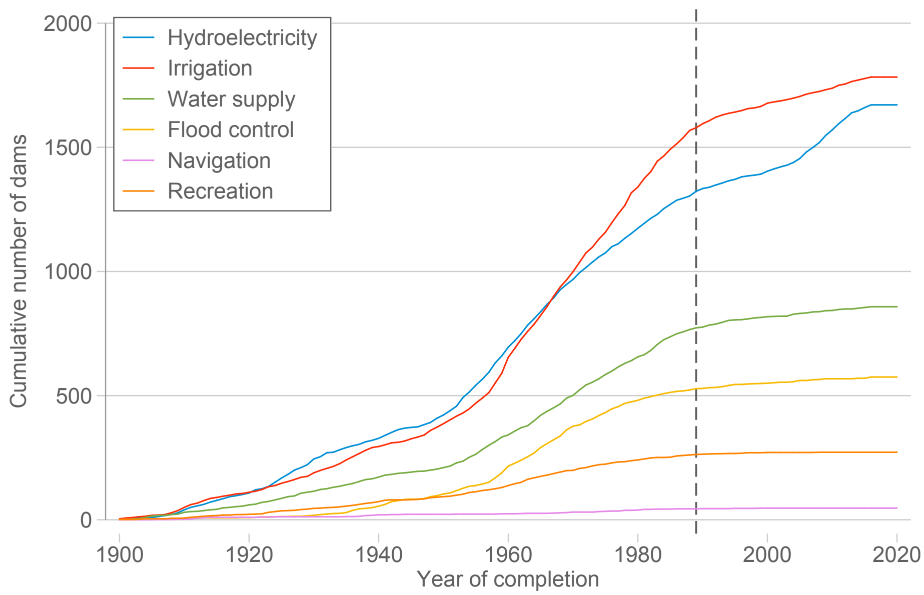
A total of 922 large dams across 77 countries have been built between 1989 and 2016, providing sufficient variation during the sample period. The majority of dams is located in Asia (553), followed by Europe (118), North America (92), South America (90), Africa (62) and Oceania (7), as shown in Figure A3 and A5. Notice, while the annual number of dams completed globally has slowed down over the years, dams have become increasingly larger in size, c.f. Figure A1. The two most common types of dams by primary use are hydroelectric (44.48 percent) and irrigation dams (25.94 percent), as depicted in Figure 1 and Figure 4. Switching the focus to the drainage basin level: recently dammed basins contain an average (median) of 1.48 (1) dams, with a dam height of 81 (73) meters and a reservoir capacity of

[12] Alternative global datasets include the World Register of Dams by the International Commission on Large Dams used by Duflo and Pande (2007), no geolocation only the nearest city is reported. Global georeferenced Database of Dams (GOOD2) and Future Hydropower Reservoirs and Dams (FHReD) are both geocoded but no sufficient attribute information available. Global Dam Tracker (GDAT) will be available in 2020 and have both geolocation and some statistics, it however appears that no size threshold will be imposed.

[13] According to the author of the data, the specified year variable usually refers to the construction completion or commissioning date. I manually checked all dams after 1988 and it turns out that the year variable almost always refers to the completion year. In cases where it does not, appropriate corrections are made.

1.31 (.22) cubic kilometers.

Figure 1: Dams by Primary Use



Notes: Dams are displayed by their primary use since 1900 with data from GRanD. The vertical dashed black line indicates the start of the sample, 1989.

Figure A4 in the Appendix depicts the spatial distribution of violence, with darker shadings indicating basins with a higher share of sample years with conflict. Conflict occurs in 123 out of 153 countries, with 43.59 percent of all conflict incidents taking place in Asia, followed by Africa with 40.61 percent. Panel A of Table 1 reports that 625, or 3.9 percent of all 15,923 drainage basins contain at least one newly constructed dam since 1989 and half of the treated basins contain more than one new dam. All treated basins (i.e. those with dams built after 1989) account for 8.9 percent of global conflict incidents in 1989 to 2016, as documented in the last column of Panel A. Extending the view to all basins with dams since 1900, similar proportions can be observed: dams are built in 15 percent of all basins, with a share of 30 percent of all conflict incidents. The first column of Panel B of Table 1 shows the unconditional probability of observing a conflict event in a basin in a given year of 3.40 percent. Observing the same probability in basins with dams is more than double, with 7.3 percent. Interestingly, this increase in conflict exposure seem to persist over time, as dams built before 1989 still show a close to 50 percent higher probability of conflict than basins without any dams. Overall, basin with dams experience on average more violence. This is of course not more than a first clue to a possible, positive association between dams and conflict, as multiple sources of potential biases prevent making any causal statements.

Table 1: Descriptive Statistics

	Basins without dams	Basins with dams		Total
		Pre-sample	Within-sample	
Panel A. General Information				
Number of basins	13,626	1,672	625	15,923
Share of basins	0.856	0.105	0.039	1
Number of dams since 1900	0	4,206	2,495	6,701
Number of dams since 1989	0	0	922	922
Share of conflict incidents	0.696	0.215	0.089	1
Panel B. Conflict and Dam Data (Basin-level)				
Pr(Conflict > 0)	0.031	0.046	0.073	0.034
	(0.173)	(0.210)	(0.259)	(0.181)
Rebels vs. State	0.021	0.029	0.056	0.024
	(0.145)	(0.169)	(0.230)	(0.152)
Rebels vs. Civilians	0.009	0.016	0.023	0.010
	(0.093)	(0.124)	(0.149)	(0.100)
State vs. Civilians	0.006	0.009	0.008	0.006
	(0.077)	(0.094)	(0.089)	(0.080)
State vs. State	0.000	0.001	0.001	0.000
	(0.019)	(0.024)	(0.039)	(0.020)
Pr(New dam completed > 0)	0.000	0.000	0.053	0.002
	(0.000)	(0.000)	(0.236)	(0.048)

Notes: The sample includes 15,923 basins for the years 1989-2016. "Pre-sample" refers to all drainage basins with at least one dam completed prior to 1989 and none thereafter. "Within-sample" indicates basins with at least one dam completed since 1989, regardless of any pre-existing dams. Panel A: values refer to the number/share within each of the four categories. Panel B: the unit of observation is a drainage basin and a year. Standard deviations in parentheses, where applicable.

5 Empirical Strategy and Main Results

5.1 Identification Challenges

The main challenge in estimating the causal effect of dams on conflict is reverse causality, as regions subject to violence may receive less investments in new infrastructure. If a government anticipates a rise in conflict in a certain region, it may reduce infrastructure spending into that regions today, because battles could severely damage or completely destroy dams and flood downstream communities. In addition, rebels groups operating in conflict-prone regions could attempt to seize strategic control over infrastructure located in conflict-prone regions (Berman et al., 2017). Therefore, dam placement might be endogenous to conflict as governments potentially prefer to allocate dams to peaceful, state-controlled regions. The existing literature has argued along similar lines for the case of irrigation dams, which tend to be placed in wealthy regions with high anticipated agricultural growth (Duflo and Pande, 2007; Strobl and Strobl, 2011). In either scenario, regressing conflict on dams would cause a downward bias in the OLS estimation. In other words, a positive causal effect of dams on conflict is likely to be underestimated in the OLS. In addition, classical measurement error could further attenuate coefficients towards zero.

5.2 Instrumental variable strategy

To address these issues, Duflo and Pande (2007) developed an instrumental variable strategy exploiting the fact that dams require a certain river gradient. The river gradient measures the change in elevation and is hence determined by natural conditions, and not man-made. The incurred costs of a dam site crucially depend on how well-suited the geographic conditions, such as the gradient are. The engineering literature defines a non-monotonic relationship between the river gradient and the suitability of a site for dam construction. A *moderate river gradient* (1.5-3 percent) ideally facilitates the command area of dams, as it allows water to run through the canal system and irrigation facilities in a controlled manner. In addition, a *steep river gradient* ($>$ 6 percent) favors hydropower facilities, allowing the maximal amount of water to flow through the turbines of the power plant in the shortest amount of time (Lipscomb, Mobarak and Barham, 2013).[14]

On the other hand, neither a flat nor an intermediately steep gradient offer optimal conditions for dams. A flat river gradient (< 1.5 percent) is generally unsuitable, because it prevents water from flowing and in addition flat landscapes usually do not provide the necessary geography for natural reservoirs. Finally, with an intermediately steep gradient (3-6 percent) the control over water flows is reduced as the water flows rather fast through the canal system, which additionally erodes the system.

The ideal drainage basin therefore features a combination of moderate and steep gradients along different sections of rivers. I exploit this non-linearity in the river gradient's suitability for dam placement and calculate the *fraction of moderate (1.5% to 3%) and steep (> 6 percent) river gradient* RG_i for each basin i. The lower bound of RG_i is zero, if none of the rivers of a basin fall within moderate or steep gradients. The upper bound of RG_i is one, if all rivers of a basin lay within moderate or steep gradients. A first look at the summary statistics reveals that basins with dams indeed have on average a higher RG_i, c.f. Table A1 in the Appendix. Figure 2 confirms a positive relationship between a higher fraction of suitable river gradient RG_i (darker shading) and dam placement (purple dots).

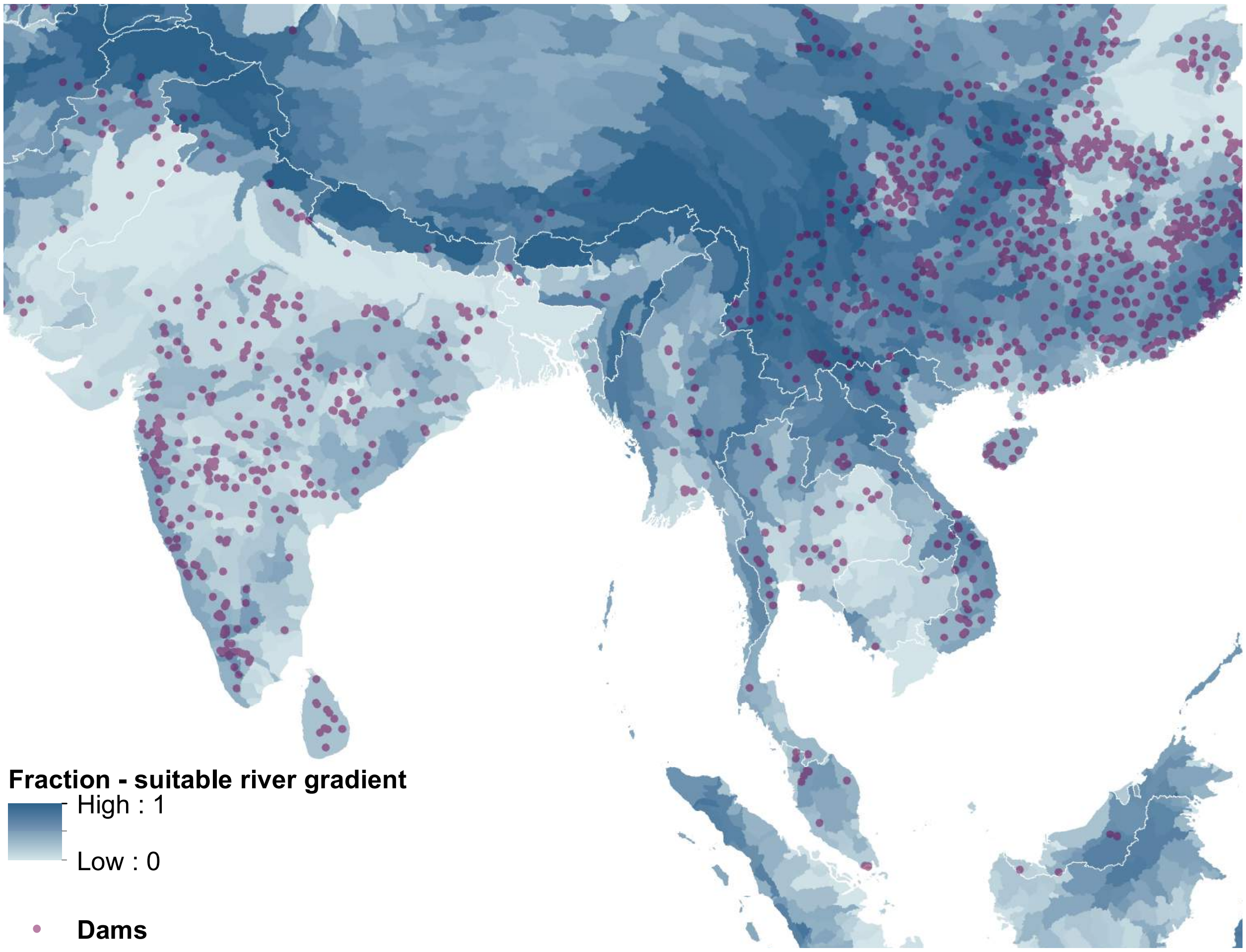
Note that RG_i is constant over time. To achieve a time-variant instrument usable for panel analysis, I interact each basin's suitable gradient share with the *total number of dams* $\overline{D}_{ct}$ in country c and year t. That gives an interacted instrument $RG_i \times \overline{D}_{ct}$, with variation across basins and years. The rationale behind the instrument follows a two-step procedure: In a first step, a government decides how many dams it wants to complete by a certain year, based on the expected national demand for e.g. electricity or agricultural goods.[15] In a second step, the government assigns the new dams to basins offering the best geographic conditions, i.e. basins with favourable river gradients. In other words, the government chooses basins with cost-efficient geographic conditions.[16] To summarize, the interaction

[14]Hydropower dams are the most common type by primary purpose in my sample since 1989 (c.f. Section 4.4). In addition, many dams have hydropower generation reported as their secondary purpose. For instance, 40 percent of dams stating irrigation as primary purpose also generate hydropower as secondary purpose.

[15]For instance, if a country completes five new dams in year t, then $\overline{D}_{ct}$ increases by five, compared to the previous period. Similar to Lipscomb, Mobarak and Barham (2013), a change in $\overline{D}_{ct}$ can best understood as a country's annual national budget for dams for that year.

[16]Duflo and Pande (2007) point out that in India the number of dams built in a year is determined on

Figure 2: First-Stage: Gradient Instrument and Dam Placement (Basin-level)



Notes: The map depicts the relationship between the suitable river gradient share in a basin and dam placement. The map zooms into parts of Asia for presentation purposes. "Fraction - suitable river gradient" is equal to RG_i, which is defined as the river share with a moderate (1.5-3%) and steep ($> 6\%$) gradient. Within a country, drainage basins with a higher share of suitable river gradient (darker shading) are favored for dam placement, whereas unsuitable regions (lighter shading) are less frequently subject to dam placements.

term allows for an empirical specification in which the number of dams per basin and year can be instrumented. The next section first outlines the empirical specification, before moving on to discuss the baseline results.

the state level and dams are then assigned to districts with suitable river gradient within a state. However, the Indian state-to-district allocation rule may not hold for all countries, as political hierarchies differ across countries. For instance, more centrist or smaller nations may decide the number of dams directly on national level. To design an instrument that can be applied to the whole world without loss of generality, I assume that the overall number of planned dams in a country is driven to some extent by a country-wide demand function for electricity or agricultural goods.

5.3 Econometric specification

Consider the following 2SLS system of equations:

$$Y_{ict} = \beta D_{ict} + \mathbf{Z}_{ict}\Gamma + \upsilon_i + \mu_{ct} + \epsilon_{ict} \tag{1}$$

$$D_{ict} = \alpha(RG_i \times \overline{D}_{ct}) + \mathbf{Z}_{ict}\Gamma + \upsilon_i + \mu_{ct} + \omega_{ict} \tag{2}$$

Equation (1) is the second stage equation with i, c and t indexing the basin, country and year, respectively. The dependent variable Y_{ict} is in most specifications a conflict dummy, which takes the value one if at least one conflict event is reported per basin and year, zero otherwise. υ_i denotes basin fixed effects and controls for time-invariant omitted variable bias. Country-year fixed effects μ_{ct} control for time-variant between-country heterogeneity. Conditional on the fixed effects structure, the estimation specification still allows for basin-year variation. D_{ict} is the endogenous variable of interest recording for each year and basin the total number of dams since 1900. Basins without dams are coded as zero. Conditional on basin fixed effects, D_{ict} only adds variation if a new dam had been built in a basin between 1989 and 2016. In other words, a basin's pre-sample stock of dams does not add any variation to D_{ict}. The empirical strategy is designed to capture the local nature of conflict and has some conceptual similarity with a difference-in-difference estimation. The coefficient of interest β estimates the basin-wide effect of a new dam on conflict, relative to untreated basins in the same country. $\mathbf{Z}_{ict}$ corresponds to a vector of geographic controls interacted with $\overline{D}_{ct}$. Note that any time-invariant control variables on their own are absorbed by the basin fixed effects.

In order to estimate the causal effect, it is necessary to instrument for the earlier discussed endogeneity in D_{ict} with the proposed river gradient instrument $RG_i \times \overline{D}_{ct}$. Equation (2) depicts the first-stage, whereby the river gradient interaction term predicts each year the number of dams in a basin. The remaining terms are defined as above.

5.4 Identifying assumptions

In order to recover the causal effect, the exclusion restriction requires the interaction term $RG_i \times \overline{D}_{ct}$ to be exogenous, conditional on controls and fixed effects. A natural concern about the exclusion restriction is that the river gradient instrument could correlate with conflict through a channel different than dam construction. Basin fixed effects account for time-invariant omitted variables, such as elevation or ruggedness. Another concern is that the time-varying part of the instrument, $\overline{D}_{ct}$, correlates with country characteristics over time, such as changes in a country's military strategy. For instance, if a government anticipates civil war in the coming years, it may start reducing infrastructure spending today. A battery of country-year fixed effects μ_{ct} control for such a potential threat.

The final concern is that the interaction term itself could violate the exclusion restriction. Since the instrument's interaction term varies by basin-year, a bias could stem from some basin-induced time variation. For instance, a country's investment in dams could trigger an inflow of international aid into the same region, targeting populations in the highlands. Elevation positively correlates with the gradient and conflict, and hence could result in a bias. I address such potential issue in two ways. First, I control for the following basin-level covariates in $\mathbf{Z}_{ict}$: average elevation, average basin gradient, river length and basin area, all interacted with $\overline{D}_{ct}$. Second and unlike existing papers, I only consider the river gradient along rivers and not across the whole basin.[17] My strategy minimizes the risk of drawing gradient information from parts of the basin far away from rivers (e.g. arid highlands). Limiting the gradient information to the river center line captures the relevant river features, without including information potentially harming the identification. The relevance of the instrument is demonstrated in the next section.

As with any instrumental variable estimate, I measure the local average treatment effect (Imbens and Angrist, 1994). The coefficient β measures the average effect among compliers, i.e. basins only containing dams because of favorable gradient conditions. Interestingly, dams built today seem to be placed in basins with a higher RG_i (c.f. first row Table A1). That could be due to a host of reasons, including stricter rules from funding bodies and better technical understanding of dams today. In any case, a higher rate of compliers appears to

[17] The gradient is calculated with 1×1 km elevation data from HydroSHEDS (Lehner and Grill, 2013).

play in the favor of this paper as it improves the external validity.

5.5 Baseline results

5.5.1 First-Stage

The predictive power of the river gradient for dam placement is tested in columns 1 to 3 of Table A2 of the Appendix. In detail, I regress a binary variable indicating whether any dam has been completed during the sample period on the area share of the different river gradient classes: 1.5-3 percent, 3-6 percent, $>$ 6 percent, with $<$ 1.5 percent being the reference group. The specification controls for country fixed effects, which means the estimations rely on within-country comparisons of basins, eliminating between-country differences. Note, columns 1 to 3 solely serve the purpose of demonstrating the relationship between river gradient and dam occurrence, but is different from (and less demanding than) the actual first stage presented in column 4. First, the outcome variable focuses on hydropower dams and results in column 1 confirm the aforementioned predictions from the engineering literature: a high fraction of moderate (1.5-3 percent) and steep ($>$ 6 percent) river gradients favor dams, both with positive and highly significant coefficients for those two categories. A steep gradient appears ideal to facilitate hydropower dams with a point estimate of 0.12, corresponding to more than a half standard deviation increase in hydropower dam occurrence. On the other hand, the suitability of a mid range river gradient (3-6 percent) is not statistically different from a flat gradient ($<$ 1.5 percent). Column 2 considers all types of dams and the results remain comparable, although the coefficient of the mid range river gradient (3-6 percent) is now statistically significant, which is likely to be driven by non-commercial dams (e.g. flood control and recreational dams). The steep river gradient maintains a positive correlation with dam occurrence, because the majority of dams in my sample produces hydropower either as primary or secondary purpose.

These results emphasize that the optimal basin offers a combination of the two favorable gradient classes. That is why the time-invariant part of the earlier discussed instrument RG_i is defined as the fraction of moderate plus steep river gradient. Results in column 3 strikingly confirm that the optimal river basin requires both, moderate (1.5-3 percent) *and* steep ($>$

6 percent) gradients, with a highly significant coefficient. The magnitude of the coefficient suggests that moving from an entirely unsuitable to a perfectly suitable river basin increases the likelihood of observing a dam by 9.36 percentage points. Finally, column 5 moves to the panel setting and estimates the first-stage of the baseline specification, with the dependent variable D_{ict} on the LHS and the instrument $RG_i \times \overline{D}_{ct}$ on the RHS. In addition, baseline controls and basin and country-year fixed effects are accounted for. The first stage regression documents a strong and positive correlation between the river gradient instrument and the number of dams in a basin and year. The F-statistic for the excluded instrument, conditional on controls, is 25.41 and hence is sufficiently strong to ensure relevance of the instrument. In other words, a bias caused by a weak instrument is highly unlikely. Section C of the Appendix demonstrates the robustness of the instrumental variable strategy. The magnitude of the highly significant coefficient can be interpreted at the average drainage basin: If a country builds a single dam in a given year, an increase in the suitable river gradient by one standard deviation (0.259) increases the probability of dam placement in that basin by 26 percent, given likelihood of observing a newly-completed dam in a basin is 0.002 (c.f. Table 1).

5.5.2 OLS, Reduced Form and 2SLS

Table 2 reports the baseline results of the OLS, reduced form and the 2SLS estimations. Standard errors allow for a spatial correlation within a 400 km radius of a basin's centroid and infinite serial correlation, for the remainder of the paper.[18] All estimations are based on the unrestricted world sample and account for baseline controls and basin and country-year fixed effects, as specified in Section 5.3. The result of the OLS is reported in column 1 of Table 2 and shows a negative association between dams and conflicts, small in magnitude and weakly significant. The reduced form in column 2 displays a positive effect of the river gradient instrument on conflict, which would be the expected direction in the case of a positive causal relationship between dams and conflict. Column 3 depicts the 2SLS results of the second stage in Equation (1), with the excluded river gradient instrument accounting for endogeneity in the number of dams per basin. The baseline coefficient measures the

[18] 400 kilometers spatial clustering is chosen, because it is the median of the minimum distances between all basins' centroids and the nearest capital.

Table 2: Baseline Results

Estimation:	OLS	Reduced F.	2SLS	2SLS	2SLS
Dependent variable:	Conflict	Conflict	Conflict	Intrastate	State vs. Rebels
	(1)	(2)	(3)	(4)	(5)
Dams	-0.007^{c}		0.032^{b}	0.031^{b}	0.029^{a}
	(0.004)		(0.013)	(0.013)	(0.011)
$RG_i \times \overline{D}_{ct}$ ($\times$ 1000 *)		0.083^{a}			
		(0.029)			
Basins	15923	15923	15923	15923	15923
Observations	445844	445844	445844	445844	445844
Kleibergen-Paap F-statistic	.	.	26.089	26.089	26.089
Mean dep. var.	.034	.034	.034	.034	.024
Basin FE	✓	✓	✓	✓	✓
Country-year FE	✓	✓	✓	✓	✓
Geographic controls	✓	✓	✓	✓	✓

Notes: The unit of observation is a drainage basin and a year. The sample includes 15,923 basins for the years 1989-2016. Column 1: OLS estimation. Column 2: reduced form estimation, with the coefficient and standard error multiplied by 1,000 (*). Columns 3-5: 2SLS estimations, with the first-stage Kleibergen-Paap F-statistics reported. The dependent variables indicate conflict incidence and are equal one if at least one conflict event occurs in a basin and year. "Conflict" considers all conflict events; "Intrastate" excludes events involving two state actors; "State vs. Rebels" only includes conflict events in which government forces fight against rebel groups. "Dams" specifies the stock of dams in a basin and year. $RG_i \times \overline{D}_{ct}$ is the instrument, where RG_i is a basin's combined river fraction with a gradient of 1.5-3% or >6%; $\overline{D}_{ct}$ is the number dams in a country and year. Geographic controls: basin size, elevation, average basin gradient and river length, interacted with $\overline{D}_{ct}$. The regressions control for basin and country-year fixed effects. Coefficients are reported with spatially clustered standard errors in parentheses, allowing for a spatial correlation within a 400 km radius of a basin's centroid and infinite serial correlation (Conley, 1999). ^{c} significant at 10%; ^{b} significant at 5%; ^{a} significant at 1%.

basin-wide change in conflict after dam completion, relative to untreated basins in the same country. The coefficient corresponds to a 3.2 percentage points increase in conflict frequency and is statistically significant at the five percent level. The direction of the effect is in line with the earlier discussed policy reports and case studies. Further, dams tend to be build in peaceful, economically prospering regions, as hypothesized earlier, implying that the effect of dams on conflict in the OLS may seem less gloomy than the actual effect. When comparing 2SLS and OLS results, the OLS indeed appears to suffer from a downward bias stemming from endogeneity in dam placement. The next two columns focus on subtypes of violence. First, intrastate violence is considered in column 4, with a dependent variable that ignores all cases in which two state forces fight each other. Column 5 restricts the dependent variable to events between government forces and rebel groups. In both cases, the coefficients are comparable in magnitude and significance to the one in column 3, suggesting that dam-induced violence may be closely related to intrastate insurgencies. Section C of the Appendix performs multiple robustness checks to ensure the validity of the instrument, including alternative specifications of RG_i and $\overline{D}_{ct}$, controlling for basin-specific time trends and river-specific variables, a cross-sectional analysis and placebo tests.

5.6 Quantification

This section quantifies the magnitude of the baseline coefficient reported in column 3 of Table 2. Evaluated at the sample mean, a new dam increases the probability of conflict in a basin from 0.034 to 0.066, which is equivalent to a 94 percent increase in frequency of conflict incidence. During the 28 year of panel data, 922 dams have been built. Studies expect a spike in the number of new dams over the next decades, and I assume a maximum (minimum) of 6727 (1143) new dams the 15 year period following 2015.[19] Provided these numbers and evaluated at the sample mean, the total of new dams constructed by 2030 is projected to increase global conflict frequency by a maximum of 5.30 (minimum of 0.90) percent. It is important to mention, that these calculations are based on the average of conflict incidents and do not consider other factors such as conflict intensity, which are discussed further below.

There are reasons to believe these predictions are rather conservative. First, dams have steadily increased in size over the years, as pointed out earlier. Larger dams and reservoirs destroy a larger share of the local economy, which I will show later, may be a key driver of violence. Second, the global supply of rivers is limited and a growing number of dams will inevitably lead to dams placed in closer proximity to each other. As a direct result, negative externalities caused by dams could multiply, as it is the case along the Mekong river where a multitude of dam facilities harm fisheries, promote land erosion and amplify the impact of

[19] An estimated 3,700 hydropower dams were planned or under construction in 2015, which provides a upper bound for this exercise (Grill et al., 2015; Zarfl et al., 2015). Of those, 17 percent, or 629 dams, were already under construction and are expected to be completed by 2030, providing a lower bound for this exercise.Reliable predictions for other types of dams are sparse. Based on the last five sample years, 55 out of 100 dams were hydropower dams. Assuming similar proportions in the future, a maximum (minimum) of 6727 (1143) can be expected over the 15 year period following the sample period. Further, I make the assumption that these new dams will be completed evenly across years, hence the unconditional probability of observing a new dam in a basin and year is 0.056 (0.010).

climate change, hence putting additional pressures on the local communities (Exchange on Environment Conflict and Cooperation, 2019; Financial Times, 2020).

6 Sensitivity Analysis

6.1 Conflict scale and intensity

To effectively mitigate the sources of violence, it is important to understand what type of violence is associated with new dams. This section addresses first the role of transboundary conflicts and then studies the severity of dam-induced violence.

6.1.1 Intrastate versus interstate conflict

Anecdotal evidence from the Nile to the Mekong suggest that dams catalyze interstate tensions. The role of interstate conflict is assessed in two ways. The first part considers only interstate conflict events, by constructing an interstate conflict indicator variable solely recording events between government forces. Direct violent confrontations between military forces have decreased in recent decades and only account for 1.2 percent of events in the sample. Results reported in column 1 of Table A3 show no evidence of an increase in conflict between government forces. Measuring interstate conflict solely based on direct confrontations between military forces ignores the possibility of indirect, state-backed violence. Proxy wars often involve third parties acting on behalf of governments, usually in return for financial and political support.

Since it seems impossible to accurately match armed groups to allied government forces directly, the second part of this section does not address interstate conflict via the dependent variable, but rather tests if dams are more likely to spark conflict in the proximity of national borders. If a dam is placed just upstream of a national border, the downstream neighbor may have an incentive to prevent or attack the dam. Columns 2 and 3 of Table A3 exclude all drainage basins shared by two nations and the coefficients are similar in magnitude to the full sample, although the standard errors are larger.[20] Interestingly, however, is that

[20]It is reassuring that splitting the drainage basins along national border did not introduce some sort of bias. An alternative to splitting drainage basins along national borders is the assignment of border basins

excluding basins within 100 km distance to national borders increases the magnitude of the coefficient, as shown in columns 4 and 5. In other words, inland dams appear more prone to conflict than those close to national borders, which further emphasizes the domestic nature of dams-related violence.

6.1.2 Conflict intensity

The indicator outcome variables seen thus far measure the occurrence, or the extensive margin of violence. Next, the conflict intensity is investigated by considering the $ln(1+x)$ of the number of events and fatalities as outcome variables. Coefficients reported in Table A4 are throughout positive and statistically significant at the five percent level. To further understand the intensity of violence, I next split conflict incidents by the number of deaths observed, in the same spirit as Nunn and Qian (2014). In detail, events are ranked by their number of associated deaths and then divided into four bins with cut-offs at the 25^{th}, 50^{th} and 75^{th} percentile, corresponding to 4, 13 and 50 fatalities per basin and year. Results reported in Table A5 and columns 1 to 3 highlight that dams primarily cause low to high intensity incidents, as shown by the coefficients, significant at the five and ten percent level. Note that the reduced magnitude of the point estimates is in line with the baseline, because dividing conflict incidents into quantiles reduces the unconditional probability of observing each intensity type to around one-fourth. Further, no statistical relationship appears to exist between dams and the most violent battles (> 50 fatalities). Overall, the coefficients' responsiveness to combats involving up to 50 fatalities is generally in line with the case study evidence and further underlines the relevance of dam-induced conflict for policy makers.

6.2 Persistence and timing of violence

To address the persistence of conflict over time, I additionally control for the lagged dependent variable in the baseline specification. Results are reported in Table A6 and show that the OLS in column 1 remains comparable to the baseline. The reduced form coefficient in colum 2 and the 2SLS coefficients in columns 3 to 5 are again significant at the five percent

to the country with the majority area basin share. Regressions following this procedure yield similar results, but run the risk of falsely assigning dams and conflict event to the wrong country.

level, although their magnitudes are reduced by one-third. The first stage remains strong and the long-run effect of an additional dam on conflict is comparable to the baseline.[21] Performing a dynamic regression may entail the risk of the Nickell Bias (Nickell, 1981). The bias in the lagged dependent variable coefficient could be substantial and amount to 0.026, hence to avoid such bias the rest of the paper does not control for the lagged dependent variable.

The remainder of this section studies the timing of conflict onset. Understanding at which stage of a dam's life cycle conflict occurs for the first time has important implications for the general understanding of the dynamics of violence and for the resource allocation of policy makers. Section 3 discusses the onset of violence along different stages of dam construction and finds - based on case study evidence - that local conflict appears more likely right after the dams are commissioned, i.e. when their economic impact fully unfolds. There are however documented cases in which an outbreak of violence is already observable towards the end of the construction period. To assess this observation quantitatively, I collected information on the construction start year for the dams in my sample, based on industry data, national and international dam registers and general internet searches. Only dams with a construction start within the sample are considered. A construction start year could be identified for 305 dams. The average (median) dam required 7.5 (5) years of construction. Regression results are depicted in Table A7. First, I rerun the baseline regression for the hand-collected subsample to achieve a benchmark and results in column 1 report a coefficient that exceeds the baseline, significant at the five percent level. Basins with dams for which no construction date could be identified are excluded from the sample. Note that so far construction years have been assigned to the pre-treatment period and the data recorded a new dam from the year of dam completion and thereafter. Column 2 repeats the exercise, with the difference that the construction years are excluded from the sample. This only changes the composition of the pool of pre-treatment observations, whereby the pool of treated units remains unaltered. The increase in the point estimate suggests that conflict might be already triggered during the construction period, as construction years appear to be on average more violent than earlier

[21] I follow Nunn and Qian (2014) and compare the full effect of a single dam (0.032) to the full effect in the dynamic regression (0.028). The latter is defined as $\tilde{\beta}/(1-\gamma)$, with the direct effect of a new dam $\tilde{\beta}$ and the indirect effect γ via the lagged dependent variable.

pre-treatment observations. Column 3 now reintroduces construction years to the sample, but reassigns them to the treatment periods. In other words, the data now records a new dam from the construction start inwards. The magnitude of the coefficient is now between the ones of column 1 and 2, which could suggest that construction years are more violent than the average pre-treatment observation, but less violent than the average treatment unit. Columns 4 to 6 repeat the exercise for intrastate conflict events and results are in line with previous columns. To summarize, these results suggest that the onset of violence might already occur during the construction year, but an escalation of violence can only be observed post completion. The next section performs synthetic controls to further investigate the timing of conflict and to verify the robustness of previous findings with an alternative identification strategy.

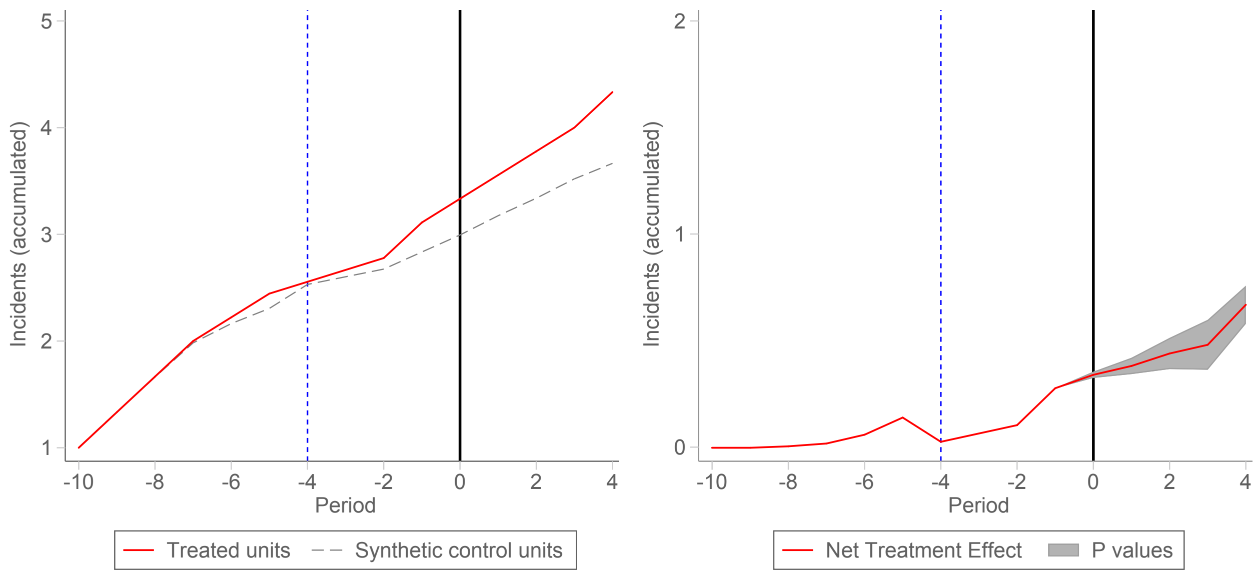
6.3 Synthetic Control Method

In recent years, synthetic controls have gained popularity among impact studies (e.g. Abadie and Gardeazabal (2003); Abadie, Diamond and Hainmueller (2010); Cavallo et al. (2013); Acemoglu et al. (2016)). The method is especially suitable for event studies with multiple treatments scattered across time, because it does not require to explicitly pre-define counterfactual observations. Instead, a data-driven process generates the counterfactual from a weighted combination of control units that matches the treated unit's pre-treatment trend. Synthetic control method may be a useful tool for policy makers who are interested in analyzing the impact of dams in sub-regions, without the need to impose parallel trend assumptions. Further, an advantage of this method over e.g. propensity score matching is that generating the counterfactual based on the pre-treated outcome variable implicitly captures all observable and unobservable determinants. The analysis in this section focuses on India, a country with one of the highest numbers of dams worldwide and subject to previous empirical work.[22] Rather than using a conflict dummy, the outcome variable is defined as

[22] I study a country rather than the whole world to keep observations in the donor pool comparable with respect to institutional characteristics, because cross-country heterogeneity may crucially affect the impact of dams on conflict, as shown later on. In addition, the world sample consists of more than 15,000 drainage basins and the literature commonly restricts the donor pool to 5,000 observations for reasons of computability. Selecting a sub-sample from the donor pool could introduce a selection bias.

the running sum of the number of conflict incidents per basin over time, which favors the convergence properties of synthetic controls (Rohner and Saia, 2019). To understand the timing of violence, an appropriate number of periods before and after the dam completion year is required. Therefore, the sample is restricted to dams built within the time window 1999 to 2010. In addition, only basins without preexising dams are considered, to avoid falsely capturing the effect of a previously constructed dam. In cases of multiple new dams in the same basin across different periods, the outcome variable is coded as missing from the year onwards the second dam is completed, following Cavallo et al. (2013).[23]

Figure 3: Synthetic Control Method for India



Notes: Synthetic Control Method for India. Dam completion is centered at period zero (black solid line) and pre-sample trends of the synthetic controls are matched to treated units up to four years prior to dam completion, i.e. left of the vertical dashed blue line. Left Panel: The red line depicts the average running sum of conflict of treated basins and the dashed gray line depicts the corresponding average of synthetic controls. Right Panel: The black line corresponds to the average net treatment effect, i.e. the difference between treated units and synthetic controls. The gray-shaded area corresponds to the probability of observing the net treatment effect by pure chance, standardized, from period -4 onward.

This gives an Indian sub-sample consisting of nine new dams and 267 potential control basins in the donor pool, with the latter being matched onto the pre-treatment running sum of conflict incidents of each treated basin to construct the counterfactual. To capture a

[23] Although unlikely, untreated basins surrounding a dam are excluded from the pool of donor observations to avoid capturing spillovers of the treatment effect. The pool of control basins is further restricted to those with a high matching quality of less than $\sqrt{3}$ times the average pre-treatment root mean squared predictive error, following Acemoglu et al. (2016).

potential surge in conflict in the years prior to dam completion, pre-trends are not matched up to the year of dam completion (black line Figure 3), but up to four years prior to completion (i.e. left of dashed blue line). The left panel of Figure 3 depicts the average running sum of conflict events of treated basins (red line) and the average of synthetic control units (gray line). The trends of treatment and controls align well during the matching period ten to four years prior to dam completion. Conflict in treated basins starts to depart from the control trend one year before dam completion and continues to increase in the years following dam completion. The right panel of Figure 3 depicts the difference between treated and control trends, with a positive net treatment effect starting just before dam completion. The graph further plots the probability of observing this increase in conflict in treated basins by pure chance, which is initially close to zero and then increases over time. Although this analysis is limited to India, the results are in line with the 2SLS results and the earlier discussed case studies arguing that the filling of reservoirs in some cases pre-dates dam completion.[24]

6.4 Spatial diffusion of violence

The analysis so far has focused on assessing the effect of dams on conflict within basins. Yet, dams transform downstream water flows and potentially affect regions beyond a treated drainage basin's boundaries. Whether to expect conflict to increase or decrease in downstream regions of dams is a priori not clear. This sections assesses the role of spatial spillovers by controlling for the upstream drainage basin.[25] Drainage basins are especially useful for this purpose, because they are constructed in a way that allows to identify adjacent basins. I follow Duflo and Pande (2007) and extend the framework of Section 5.3 by instrumenting for the number of dams in the upstream and downstream basin:

$$Y_{ict} = \beta_1 D_{ict} + \beta_2 D^U_{ict} + \mathbf{Z}_{ict}\Gamma_1 + \mathbf{Z}^U_{ict}\Gamma_2 + \upsilon_i + \mu_{ct} + \epsilon_{ict} \tag{3}$$

$$\Delta_{ict} = \alpha_1(RG_i \times \overline{D}_{ct}) + \alpha_2(RG^U_i \times \overline{D}^U_{ct}) + \mathbf{Z}_{ict}\Gamma_1 + \mathbf{Z}^U_{ict}\Gamma_2 + \upsilon_i + \mu_{ct} + \omega_{ict} \tag{4}$$

[24]The data collected for Section 6.2 include information on the reservoir filling year (where available). In some cases the reservoir is started to be filled already one to two years prior to dam completion.

[25]This section follows the existing literature and focuses on the impact of dams located within and upstream of basins, whereas dams located downstream of basins have widely been found to have no effect (Duflo and Pande, 2007; Strobl and Strobl, 2011).

Equation (3) appends the baseline second stage equation with the number of dams in the upstream basin D^U_{ict} and a vector of controls in the upstream basin $\mathbf{Z}^U_{ict}$. Notably, β_1 measures the effect of an additional dam on conflict in basin i and β_2 measures the effect of an additional dam in upstream basin j on conflict in basin i. Recall, the spatial unit is a basin, hence β_1 and β_2 measure the effect of different dams on a single basin and not the effect of a single dam on different basins. Equation (4) signifies the first-stage equation, with the number of dams in the own basin D_{ict} and in the upstream basin D^U_{ict}, both represented by Δ_{ict}. In addition, controls for the upstream basin are added to maintain the same structure of controls throughout the 2SLS system of equations.

Before performing the two instrument version, I first estimate the isolated effect of an additional upstream dam on conflict in my own basin, by omitting D_{ict}, $\mathbf{Z}_{ict}$ and $RG_i \times \overline{D}_{ct}$ from the econometric specification. This is essentially the upstream version of the baseline. Results are depicted in column 1 of Table A8 in the Appendix. The coefficient of an additional dam in the upstream basin on conflict in the own basin is positive, insignificant and the instrument is weak. The latter may also be due to a reduced sample size, because 20 percent of the basins are geographically isolated and do not have an upstream basin.[26] Next, I replicate the baseline regression for the sub-sample at hand, to achieve a benchmark estimate. Results in column 2 show a coefficient significant at the five percent level and which exceeds the baseline coefficient of Table 2 by nine percentage points. The next column additionally controls - without instrumenting - for the number of dams in the upstream basin. Simply controlling for upstream dams does not however recover the causal effect, because just like the main basin, upstream basins could suffer from endogenous dam placement. Therefore, column 4 instruments now for both, dams in your own basins and upstream. The main effect in the first row is in line with the benchmark in column 2 and the coefficient of an additional upstream dam in the second row is close to zero and insignificant. The reported Kleibergen-Paap F-statistic decreases to 7.05, because predicting $\widehat{D_{ict}}$ and $\widehat{D^U_{ict}}$ from the own and the upstream basin's river gradient is more demanding than the baseline first stage. The regression reported in column 5 additionally accounts for upstream basin controls $\mathbf{Z}^U_{ict}$,

[26] In addition, tributaries and river head basins are coded as zero, as these units neither have upstream basins.

hence estimates the second stage as specified in Equation (3). Including more controls further reduces the predictive power of the instrument and the Kleibergen-Paap F-statistic drop far below ten, hence prohibits a causal interpretation of the results.

To summarize, results in columns 2 to 4 show no consistent evidence for spillovers in conflict from upstream basin. Further, the baseline effect of a new dam in the own basin appears widely unaffected by dams in the upstream basin, the very weak instrument of the specification in column 5 does however not allow to dismiss spatial spillovers with certainty. To further assess whether the main effect in my own basin is potentially influenced by upstream dams, I exclude jointly-determined basins from the regression, i.e. basins with new dams that at least one new dam in the upstream basin. Excluding these basins allows to rule out a potential bias in β_1 stemming from dams in the upstream basin. The according coefficient in column 6 is consistent with and even exceeds the potentially biased result in column 2. Whether the increase in magnitude is due to sample selection or to conflict-mitigating spillovers remains unclear. In either case, column 6 accounts for a potential bias in the main coefficient and finds a positive association between dams and conflict, in line with the baseline.

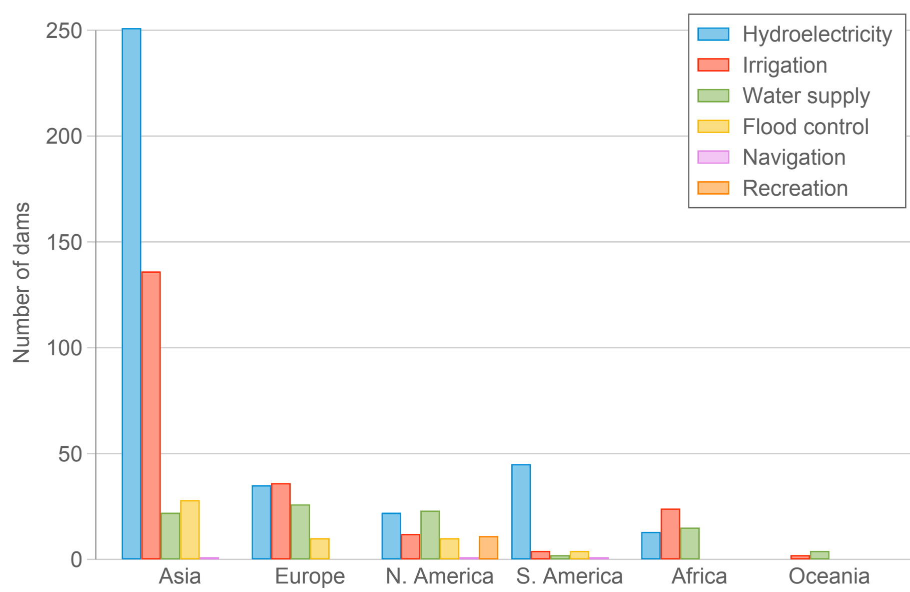
6.5 Extensive margin of dams and main dam types

Half of the treated basins receive more than one new dam between 1989 and 2016. To test whether the first dam causes more conflict than consequent ones, I exclude consequent dams during the sample period by coding D_{ict} as missing from the year onward a basin receives a second dam.[27] Results in columns 1 to 3 of Table A9 show that the first dam during the sample period is especially harmful with a point estimate about twice the magnitude of the baseline. The first dam causing more damage is indeed plausible, because by the time a second dam is completed, a significant part of the local population has already been displaced. Note that D_{ict} now only adds variation to a basin's pre-sample stock of dams via the first finished dam during the sample period. The ideal regression would further exclude

[27]Furthermore, to measure the effect of the single first sample dam on conflict, basins with several dams completed in a single year are excluded from the sample; these basins only account for three percent of all treated basins in this sub-sample and including them yields similar results.

all pre-sample dams from the analysis to measure the true extensive margin. This would however result in a binary D_{ict} and hence be estimated inconsistently. What can be done is to return to the unrestricted sample and now exclude basins with pre-sample dams, to test if past experience in dam-building reduces the probability of conflict. Columns 4 to 6 of Table A9 do exactly this and results exceed the baseline, but are smaller than those in columns 1 to 3. Overall, the results suggests that the first dam causes more local conflict than consequent ones.

Figure 4: Dams by primary use since 1989



Notes: Dams are displayed by their primary use and continent. Only dams completed within the sample period (1989-2016) are considered. Dams without information on primary use (about 20%) are omitted. Continents are ranked by total of completed dams.

Next, dams are studied along the purpose they serve. The impact on the surrounding environment and economy can vastly vary by dam type. Especially dam types reducing the local water supply have been found to impair the local economy and therefore have the potential to cause more conflict (World Commission on Dams, 2000). I rerun the baseline specification with the difference that now D_{ict} separately considers the three most common dams by primary and secondary use, namely hydropower, irrigation and water supply dams.[28] Results are reported in Table A10. All coefficients document a positive association

[28]Both, primary and secondary dam purposes are considered here, to obtain sufficient variation in the

between dam placement and conflict, with hydropower dams being less harmful than irrigation and water supply dams. Coefficients are less precisely estimated and just miss the ten percent significance threshold, likely to be due to the reduced variation in D_{ict}. Differences in coefficient magnitude across types are in line with case studies arguing that the local impact of hydropower dams is comparatively small, as water is stored and released according to the electricity demand, but none of the water is removed from the river system. Contrarily, water supply dams frequently shift water to other locations such as faraway industrial hubs and irrigation dams remove water from the river system to nearby irrigation facilities (World Commission on Dams, 2000).

7 Country Heterogeneous Effects

This section discusses the sensitivity of dam-related violence to several country characteristics that are frequently associated with conflict. Country heterogeneity is measured for the year 1988, i.e. one year prior to the sample start. Using constant pre-sample values rather than time-variant panel data may introduce some measurement error, but avoids the greater threat of reverse causality. In other words, conflict could potentially alter a country's institutional and economic conditions, hence directly affect the tested heterogeneity. I define an indicator variable I_c that takes the value one if a country has an above-median score in the tested heterogeneity. As measured of 1988 only, the potential heterogeneity enters the equation only once and is measured via β_2:

$$Y_{ict} = \beta_1 D_{ict} + \beta_2 (D_{ict} \times I_c) + \mathbf{Z}_{ict}\Gamma + \upsilon_i + \mu_{ct} + \epsilon_{ict} \quad (5)$$

Where $D_{ict} \times I_c$ indicates dam in the sub-sample of countries subject to the tested heterogeneity. All other variable definitions remain unchanged from Equation (1) and the stand-alone term I_c is absorbed by the fixed effects. Note, the endogenous variable D_{ict} occurs now twice in the second-stage equation and to achieve a just-identified 2SLS system, the instrument is

number of newly-completed dams of each sub-type.

additionally interacted with I_c in the first-stage equation:

$$\Delta_{ict} = \alpha_1(RG_i \times \overline{D}_{ct}) + \alpha_2(RG_i \times \overline{D}_{ct} \times I_c) + \mathbf{Z}_{ict}\Gamma + \upsilon_i + \mu_{ct} + \omega_{ict} \qquad (6)$$

Where Δ_{ict} represents the two instrumented terms D_{ict} and $D_{ict} \times I_c$. The regressor I_c and interaction terms $RG_i \times I_c$ and $\overline{D}_{ct} \times I_c$ are absorbed by the fixed effects structure and hence omitted from the first-stage equation.

Table A11 first considers heterogeneity across the level of democratization with data from the Polity 4 Project (Marshall, Gurr and Jaggers, 2012). The estimates in column 1 document that highly democratic countries are not statistically differently exposed to conflict than less democratic countries.[29] Next, I test the role of country-wide differences in the level of ethno-linguistic polarization. Polarization is an established index to approximate inter-group division and is defined according to Montalvo and Reynal-Querol (2005) and based on data from the Ethnologue (Lewis, Simons and Fennig, 2018). Results in column 2 show a positive, although statistically insignificant coefficient of the interaction term, suggesting that countries with an above-median polarization appear similarily exposed to violence as less polarized countries.[30] Column 3 investigates the role of real income per capita with data from the Penn World Table (Feenstra, Inklaar and Timmer, 2015). Results document very weak evidence for a lower conflict exposure of richer nations, with a coefficient of the interaction term statistically different from zero just above the ten percent level. Next, the role of the agricultural sector is addressed in more detail. As outlined earlier, dams affect the local endowment of arable land and water, both essential input factors for farming. Column 4 tests heterogeneity across countries' agricultural endowment with data by Faostat (Food and Agriculture Organization, United Nations, 2013). The highly significant coefficient of the interaction term documents that dams are especially harmful in countries with above-median agricultural land cover shares, reflected in a coefficient with a magnitude of about twice the

[29] The democracy variable is a compound of four dimensions: competitiveness of executive recruitment, openness of executive recruitment, constraint on chief executive and competitiveness of political participation.

[30] A subnational analysis of ethnic differences and details on the data construction can be found in Section 8.2. The ethnic fractionalization index (1 - Herfindahl-Hirschman index) measuring ethnic diversity yields similar results. The ethnic segregation index measuring the spatial division of groups, as defined in Alesina and Zhuravskaya (2011), yields inconsistent results across different data sources.

one of the baseline. In turn, countries with a low dependency on the agricultural sector show no sign of an elevated exposure to conflict, with a first row coefficient statistically indifferent from zero. Lastly, to connect agriculture to income, I test whether countries generating a high share of income from agriculture are more prone to conflict, with data from the World Development Indicators (World Bank, 2019). The correlation coefficient between real GDP per capital and share of agricultural income is -.55 and statistically significant at the 1 percent level. The regression coefficient reported in column 5 is positive, significant at the ten percent level and of similar magnitude as the previous column. While these findings point towards the central role of agriculture for dam-related violence, country-level heterogeneous effects could potentially be confounded by many other institutional characteristics. The next section addresses this shortcoming, by investigating the economic impact of dams at the local level.

8 Channels and Mechanisms

Thus far, the analysis has focused on the nature of dam-induced violence, whereas the underlying economic and institutional conditions associated with an outbreak of conflict have not yet been discussed. Section D of the Appendix investigates the local economic impact of dams, with disaggregated GDP data, satellite night-time lights and remotely-sensed data measuring the impact on agricultural productivity. The results are generally in line with those of Duflo and Pande (2007) who find that Indian irrigation dams are a double-edged sword with regard to their local economic impact. On the one hand, poverty increases as dams incisively disrupt economic activity in their vicinity. On the other hand, the remaining agricultural land may become more productive, as farmers gain access to better irrigation technologies. Although the direction of effects are equivalent, the coefficients in Section D are however less precisely estimated. This could be due to a host of reasons, including that some countries are more prone to adverse local shocks than others (c.f. sections 7 and 8) and that the economic impact might differ across dam types.[31]

[31] Results considering the economic impact of sub-types of dams are not shown, because of weak first-stage F-statistics.

8.1 Weak states: constrained political competition

This section spots light on institutional features likely to affect the quality of public policies and the implementation of dams. The qualitative work discussed in previous sections frequently argues that governments' reluctance to account for the preferences of affected villagers located in the dammed regions is a common element of disputes. Put in a more structured way, once the decision is made to build a dam, the central planner has to solve a bargaining process with local communities over the compensation of economic losses.[32] This bargaining process may be more likely to fail if a lack of *political competition* allows the dominant political regime to suppress or ignore the preferences of local interest groups. Besley, Persson and Sturm (2010) explore the theoretic and empirical implications of political competition and find, for US states, that a low degree of electoral competition results in non-optimal, growth-restricting policies.

To empirically investigate whether these theoretic considerations are reflected in the data, I investigate facets of political competition below. In detail, I first test whether countries with constrained freedom to express oppositional policy preferences results in a higher exposure to conflict. In detail, I construct a variable that indicates countries in which the political arena limits oppositional views and alternative political platforms, based on pre-sample data from the Polity 4 Project (Marshall, Gurr and Jaggers, 2012). Columns 1 and 2 of Table A12 provide a balancing test showing that the instrument appears to have no systematic association with variables measuring political competition. The main result is reported in column 1 in Panel A of Table 3 and shows weak evidence for elevated levels of conflict in countries with restricted competitiveness of participation, with a coefficient significant at the ten percent level and exceeding the baseline in magnitude. Column 3 further tests whether the negative economic impact of dams is rooted in the sub-sample of countries with low political competition, a weak first stage however prevents to derive

[32] In a vast majority of the reviewed case studies, local opposition is unsuccessful in reversing underway dams projects. One rare example is the Myitsone dam in Myanmar, where violent clashes between government and rebel forces lead to halt of the dam construction. When the dam construction started in 2009, anti-state grievances between the ethnic Kachin Independence Army and government forces exacerbated. Looming fears among the affected communities, mostly minorities, over agricultural land losses, flooding and earthquakes translated into an escalation of violence, marking the end of the seventeen year ceasefire between the rebels and the state. As a consequence, an estimated 20,000 people had been displaced (Exchange on Environment Conflict and Cooperation, 2019; ICTA-UAB, 2019).

reliable conclusions.

I repeat the exercise with a second index measuring the electoral competitiveness with data from DPI (Beck et al., 2001). In this variable, dictatorships with unelected legislatures receive a low score, whereas highly competitive political systems relying on coalition building receive a high score. Results are reported in the even columns of Panel A and confirm the aforementioned findings. This may suggests that governments may be less devoted to the demands of affected groups, because elections are unfair and hence repercussions at the next election are unlikely.

8.2 Ethnic grievances

Previous sections showed that wealth generated from dams may be unevenly distributed within countries, with the directly affected communities being more likely to suffer economic losses. According to the case study literature, spatial inequalities can be amplified when compensation schemes lack behind, as it often is the case when ethnic minorities are among the displaced populations (World Commission on Dams, 2000). This observation is in line with the quantitative literature that generally finds a negative association between ethnic fragmentation and public good provision.[33] This section takes a granular view at the role of ethnic grievances and tests how local ethnic power relations affect the local impact of dams. In detail, I calculate two established indices describing the ethnic composition at the drainage basin level. The first, ethnic polarization, captures the degree of division of ethnic groups and further has been found to correlate with conflict. The second, ethnic fractionalization, measures the degree of ethnic diversity (Montalvo and Reynal-Querol, 2005). To construct indices at the basin level, I populate the language map of the Ethnologue's World Language Mapping System with data from the Global Human Settlement Layer for the year 1975 (Lewis, Simons and Fennig, 2018; Pesaresi et al., 2019), following the procedure prionereed in Desmet, Gomes and Ortuño-Ortín (2019) and refined in Eberle et al. (2020). This results in a 1×1 kilometer gridded dataset with information on the number of speakers of each

[33] For an overview of the literature on the link between ethnic diversity and public goods, please consult Desmet, Ortuño-Ortín and Wacziarg (2012).

Table 3: Channel Analysis: Political Competition and Ethnic Grievances

Estimation:	2SLS			
Dependent variable:	Conflict		ln(GDP/area)	
	(1)	(2)	(3)	(4)
Panel A. Constrained political competition				
Dams	-0.031	-0.024	-0.407	-0.357
	(0.034)	(0.036)	(1.015)	(0.887)
Dams × Low participation comp.	0.073^c		-0.737	
	(0.038)		(1.193)	
Dams × Low elect. competitiveness		0.066^c		-0.811
		(0.039)		(1.078)
Basin	13280	13280	1793	1793
Observations	371840	371840	50216	50216
Kleibergen-Paap F-statistic	11.256	12.529	2.270	2.958
Panel B. Ethnic grievances				
Dams	0.015	0.020	-0.656	-0.745
	(0.013)	(0.014)	(0.534)	(0.581)
Dams × High Polarization (basin)	0.022^a		-0.452^c	
	(0.008)		(0.269)	
Dams × High Fractionalization (basin)		0.022^a		-0.420
		(0.008)		(0.281)
Basins	12225	12225	1605	1605
Observations	342300	342300	44956	44956
Kleibergen-Paap F-statistic	8.513	8.259	8.197	6.936
Basin FE / Country-year FE / Geographic controls	✓	✓	✓	✓

Notes: 2SLS estimations. The unit of observation is a drainage basin and a year. The sample sizes vary with the data availability of the dependent variable. Columns 1-2: The dependent variable indicates conflict incidence and is equal one if at least one conflict event occurs in a basin and year. Columns 3-4: The dependent variable measures income per km^2 with "Gross Cell Products" data from G-Econ, available for 1990, 1995, 2000 and 2005. Panel A: "Low participation comp." indicates countries with a below-median score in competitiveness of participation for the year 1989. "Low elec. competitiveness" indicates countries with a below-median score in executive electoral competitiveness in 1989. Panel B: "High polarization (basin)" indicates basins with an above-median ethno-linguistic polarization within a country in 1975. "High fractionalization (basin)" indicates basins with an above-median ethno-linguistic fractionalization within a country for the year 1975. Geographic controls: basin size, elevation, average basin gradient and river length, interacted with $\overline{D}_{ct}$. The regressions control for basin and country-year fixed effects. Coefficients are reported with spatially clustered standard errors in parentheses, allowing for a spatial correlation within a 400 km radius of a basin's centroid and infinite serial correlation (Conley, 1999). ^{c} significant at 10%; ^{b} significant at 5%; ^{a} significant at 1%.

language group worldwide.[34] I then construct a binary variable taking a value of one if

[34] One useful feature of language maps is the global availability of the data. In addition, each language has a language tree that describes its affiliation to broader language families. I follow Desmet, Ortuño-Ortín and Wacziarg (2012) and calculate ethnicity indices at the different subdivisions along language trees. For the regression analysis, I choose the second highest aggregation level, because it has been shown that higher

a basin has an above-median level of polarization *within a country*. About half of the dams are located in regions with an above-median level polarization. Columns 3 and 4 of Table A12 provide a balancing analysis for the tested variables and report no statistically relevant association with the river gradient instrument. The empirical specification allowing for basin-level heterogeneity is as follows, with all variables as defined above:[35]

$$Y_{ict} = \beta_1 D_{ict} + \beta_2(D_{ict} \times I_i) + \mathbf{Z}_{ict}\Gamma + \upsilon_i + \mu_{ct} + \epsilon_{ict} \tag{7}$$

$$\Delta_{ict} = \alpha_1(RG_i \times \overline{D}_{ct}) + \alpha_2(\overline{D}_{ct} \times I_i) + \alpha_3(RG_i \times \overline{D}_{ct} \times I_i) + \mathbf{Z}_{ict}\Gamma + \upsilon_i + \mu_{ct} + \omega_{ict} \tag{8}$$

The first two columns in Panel B of Table 3 show that dams built in regions with a relative high degree of ethnic polarization and fractionalization substantially amplify conflict. Further, regions with high polarization and fractionalization experience larger negative shocks to the local economy, as shown in columns 3 and 4. These results are in line with the hypothesis that complex ethnic ties are linked to a lower provision of public goods (e.g. in this case a lack of compensation schemes). Consequently, higher economic losses endured by affected groups may in turn further reduce the opportunity cost of fighting, hence could explain the surge in violence.

9 Conflict-Mitigating Institutions and Policies

The first part of this section studies institutional features arguably strengthening the bargaining power of affected communities. One key feature to keep a government's action in check is an independent, unbiased judicial system. Therefore, the role of the judicial quality is tested in column 1 of Table 4 with pre-sample data from the Political Constraints Database (Henisz, 2002). The results show that countries with an above-median degree of judicial independence are on average less likely to experience conflict. Next, I test the role of

aggregation levels are relevant to capture long-standing cleavages between groups that lead to conflict. The second highest level is preferred over the highest level, to achieve a more variation at the local level.

[35]Following Nunn and Qian (2014), the double interaction containing the shift part of the instrument does not enter $\mathbf{Z}_{ict}$. Moving $(\overline{D}_{ct} \times I_i)$ to $\mathbf{Z}_{ict}$ yields similar coefficients, although of lower significance.

property rights with data from the Economic Freedom of the World Dataset Lawson, Murphy and Williamson (2016). Secure property rights may constitute hurdles to governments to nationalize private property in the cases where land owners oppose controversial projects. Results in column 2 of Table 4 show no sign of a reduction in violence in countries with protected property rights.

Table 4: Policy Analysis: Institutional Quality and Transboundary Water Treaties

Estimation:	2SLS	2SLS	2SLS	2SLS
Dependent variable:	Conflict	Conflict	Conflict	Conflict
	(1)	(2)	(3)	(4)
Dams	0.039^{a}	0.038	0.039^{a}	0.043^{a}
	(0.013)	(0.039)	(0.012)	(0.016)
Dams × Independent judiciary	-0.041^{b}			
	(0.019)			
Dams × Property rights		-0.003		
		(0.044)		
Dams × Treaty			-0.044^{a}	-0.036^{c}
			(0.017)	(0.020)
Basins	13505	11436	15923	13924
Observations	378140	320208	445844	389872
Kleibergen-Paap F-statistic	9.269	11.66	10.974	7.975
Mean dep. var.	.037	.029	.034	.034
Mean dep. var. - interaction group	.004	.007	.044	.044
Basin FE	✓	✓	✓	✓
Country-year FE	✓	✓	✓	✓
Geographic controls	✓	✓	✓	✓

Notes: 2SLS estimations. The unit of observation is a drainage basin and a year. The dependent variable indicates conflict incidence and is equal one if at least one conflict event occurs in a basin and year. Columns 1-2: Countries included in the sample depend on data availability of the tested heterogeneity. Column 1: "Independent judiciary" indicates countries with an independent judiciary branch in 1988, with data from the Political Constraints Database. Column 2: "Property rights" indicates countries with an above-median level in the legal structure and the security of property rights in 1988, with data from the Economic Freedom of the World Dataset. Column 3: "Treaty" indicates basins subject to a transboundary water treaty in the pre-sample period. Column 4: Further excludes basins with dams in the pre-sample period. Geographic controls: basin size, elevation, average basin gradient and river length, interacted with $\overline{D}_{ct}$. The regressions control for basin and country-year fixed effects. Coefficients are reported with spatially clustered standard errors in parentheses, allowing for a spatial correlation within a 400 km radius of a basin's centroid and infinite serial correlation (Conley, 1999). ^{c} significant at 10%; ^{b} significant at 5%; ^{a} significant at 1%.

These results provide important insights into institutional aspects relevant for conflict prevention. While interesting, this type of analysis however does not inform policy in an effective way, because improving the institutional quality is a long and slow process relying to a vast extent on a government's willingness to induce change. The second part of this section therefore focuses on the role of funding bodies. Like most large-scale infrastructure projects, dams are commonly funded by development banks and other international funding bodies. Some creditors, including the World Bank, require dam-building nations to sign transboundary water treaties as prerequisite to receiving funding (Strobl and Strobl, 2011). While treaties are primarily designed to prevent transnational disputes between neighboring nations, they frequently account for local elements, such as a monitoring of the environmental impact and conflict resolution mechanisms. For instance, more than half of the treaties signed prior to 1989 account for conflict resolution mechanisms, including the formation of a permanent judicial organ or a diplomatic channel (Wolf, 1999). To test the effectiveness of transboundary treaties as conflict-mitigating tool, I employ data from the International Freshwater Treaties Database and construct a dichotomous variable indicating if a basin had been subject to a treaty before the sample start in 1989.[36] Worldwide 41 percent of all basin have been subject to a treaty in the past. Treaties usually cover large geographies within and across countries, such as entire river basins. The econometric specification is identical to Section 8.2, which allows for within-country variation. Since treaty information is solely based on the pre-sample period, the variable captures the persisting effect of a treaty on a region over time. Results are reported in column 3 of Table 4 and show that basins subject to treaties in the past experience substantially less conflict today. One might suspect that previous experience in dam building may drive the negative coefficient of the interaction term. To account for this, column 4 excludes all basins that received a dam prior to the sample start. Although the coefficient of the interaction term is now smaller in magnitude and significant only at the ten percent level, the result suggests that past dam-building expertise may not be the main conflict-mitigating factor, but rather the state capacity imposed by treaties.

[36]Conflicted regions may be more likely sign treaties, therefore relying on pre-sample values instead of a panel version prevents to introduce endogeneity.

10 Conclusion

Dams play an important role in satisfying the world's increasing demand for electricity and food and in the management of erratic climate conditions such as floods and droughts. However, critics have argued that dams come at the cost of spatially dispersing wealth and frequently causing conflict in their vicinity. The ongoing trend towards larger sized dams necessitates a better understanding of the local impact of dams, because larger dams are likely to put even more pressure on communities directly affected by their construction.

This paper is the first systematic attempt to understand the link between infrastructure and violence across the world, with the intention to initiate a discussion on how to build conflict-free infrastructure in the future. The disaggregated nature of the empirical strategy not only allows to precisely pin down the local character of dam-induced violence, but further enables to test relevant channels and policies relevant for other types of infrastructure. I find strong and robust evidence for an increase in conflict in regions with new dams. While dams on average increase local violence, it does not necessarily imply all dams are harmful. Violence is conditional on institutional characteristics and the ethnic fabric in the affected basins. The analysis suggests that a lack of political competition may result in poorly planned dams that are failing to account for the economic preferences of directly affected communities.

One powerful stimulus to build better dams are transboundary water treaties. They are an effective tool to transparently plan and implement dams and other types of infrastructure and to educate governments, especially in countries without the necessary state capacity to mediate potential disputes. It is the responsibility of funding organizations to keep governments accountable for their actions, by making credits conditional on the implementation of conflict-preventive measures and the assessment of local impact throughout all planning and construction phases of dams.

This study lays the foundation for future research on the link between infrastructure and conflict. The first promising avenue of future research is to carefully study smaller-sized and subtypes of dams, to gain a better understanding of the type-specific impact, as well as the underlying drivers of endogeneity in OLS estimations. The second is to study the relationship of dams with climate change, public health and the environmental impact, as these aspects are not only important research topics on their own, but further could contribute to the understanding of the causes of conflict. Finally, the assessment of other types of infrastructure and civil unrest is needed, in order to promote awareness of a systematic problem that requires governments to go beyond the paradigm of the cost-benefit analysis. In conclusion, this study constitutes the first steps towards quantifying the human toll of infrastructure, and more attention is required to build conflict-free infrastructure in the future.

Appendix

A Tables

Table A1: Descriptive Statistics, River Gradient Instrument (Basin-level)

	Basins without dams	Basins with dams		Total
		Pre-sample	Within-sample	
Basin fraction - River gradient IV (RG_i)	0.339	0.462	0.607	0.362
	(0.273)	(0.230)	(0.217)	(0.274)
Basin fraction - Gradient $< 1.5\%$	0.549	0.379	0.223	0.519
	(0.347)	(0.285)	(0.229)	(0.346)
Basin fraction - Gradient $1.5 - 3\%$	0.159	0.186	0.143	0.161
	(0.134)	(0.115)	(0.099)	(0.131)
Basin fraction - Gradient $3 - 6\%$	0.112	0.159	0.170	0.119
	(0.117)	(0.104)	(0.098)	(0.117)
Basin fraction - Gradient $> 6\%$	0.179	0.276	0.464	0.201
	(0.275)	(0.275)	(0.286)	(0.282)

Notes: The sample includes 15,923 basins. "Pre-sample" refers to all drainage basins with at least one dam completed prior to 1989 and none thereafter. "Within-sample" indicates basins with at least one dam completed since 1989, regardless of any pre-existing dams. The unit of observation is a drainage basin, as all variables are time-invariant. The first variable "Basin fraction - River gradient IV (RG_i)" is the suitable basin share for dams RG_i, which is defined as a basin's river fraction with a gradient of 1.5-3% or >6%. The remaining variables summarize the share of each gradient category along rivers. Standard deviations in parentheses.

Table A2: First-Stage Estimates

	(1) Any Hydrop. Dam	(2) Any Dam	(3) Any Dam	(4) Dams
Gradient $1.5-3\%$	0.0252^{a}	0.0241^{b}		
	(0.0077)	(0.0107)		
Gradient $3-6\%$	0.0040	0.0659^{a}		
	(0.0141)	(0.0185)		
Gradient $>6\%$	0.1203^{a}	0.1327^{a}		
	(0.0174)	(0.0201)		
RG_i			0.0936^{a}	
			(0.0103)	
$RG_i \times \overline{D}_{ct}$				0.0026^{a}
				(0.0005)
Basins	15923	15923	15923	15923
Observations	15923	15923	15923	445844
F-statistic	38.016	52.562	72.071	25.413
Data	Cross-section	Cross-section	Cross-section	Panel
Geographic controls	✓	✓	✓	✓
Country FE	✓	✓	✓	✓
Basin FE				✓
Country-year FE				✓

Notes: OLS estimations. The sample includes 15,923 drainage basins. Columns 1-3: the unit of observation is a drainage basin (cross-section). The dependent variables indicate if at least one dam has been completed in a basin between 1989-2016. The dependent variable of column 1 is restricted to hydropower dams (by main and secondary purpose). The independent variables measure the river fraction within each gradient category. RG_i is the time-invariant part of the instrument and measures a basin's combined river fraction with a gradient of 1.5-3% or >6%. Geographic controls: basin size, elevation, average basin gradient and river length. The regressions control for country fixed effects. Column 4: the unit of observation is a drainage basin and a year for 1989-2016 (panel). The dependent variable reports for each basin and year the stock of dams. $RG_i \times \overline{D}_{ct}$ is the instrument, with $\overline{D}_{ct}$ measuring the number dams in a country and year. Geographic controls: basin size, elevation, average basin gradient and river length, interacted with $\overline{D}_{ct}$. The regression controls for basin and country-year fixed effects. Coefficients are reported with standard errors clustered at the drainage basin level in parentheses. ^{c} significant at 10%; ^{b} significant at 5%; ^{a} significant at 1%.

Table A3: Interstate Conflict and Border Regions

Estimation:	2SLS	2SLS	2SLS	2SLS	2SLS
Dependent variable:	Interstate	Conflict	Intrastate	Conflict	Intrastate
	(1)	(2)	(3)	(4)	(5)
Dams	0.001	0.028^b	0.028^b	0.037^b	0.036^b
	(0.001)	(0.013)	(0.013)	(0.016)	(0.016)
Basins	15923	13473	13473	10841	10841
Observations	445844	377244	377244	303548	303548
Mean dep. var.	0	.032	.032	.032	.032
Kleibergen-Paap F-statistic	26.089	18.645	18.645	12.917	12.917
Basin FE	✓	✓	✓	✓	✓
Country-year FE	✓	✓	✓	✓	✓
Geographic controls	✓	✓	✓	✓	✓
Exclude border basins		✓	✓	✓	✓
Exclude basin within 100km to borders				✓	✓

Notes: 2SLS estimations. The unit of observation is a drainage basin and a year. Column 1: Only interstate conflict incidents are considered. Columns 2 and 3: Drainage basins ranging across country borders are excluded. Columns 4 and 5: Basins with the centroids located within 100 km of national borders are excluded. The dependent variables indicate conflict incidence and are equal one if at least one conflict event occurs in a basin and year. "Interstate" considers only conflict events between government military forces; "Conflict" considers all conflict events; "Intrastate" excludes events involving two state actors. "Dams" specifies the stock of dams in a basin and year. Geographic controls: elevation, gradient, river length and basin size, interacted with $\overline{D}_{ct}$. The regressions control for basin and country-year fixed effects. Coefficients are reported with spatially clustered standard errors in parentheses, allowing for a spatial correlation within a 400 km radius of a basin's centroid and infinite serial correlation (Conley, 1999). ^{c} significant at 10%; ^{b} significant at 5%; ^{a} significant at 1%.

Table A4: Conflict Intensity: The Intensive Margin of Violence

Estimation:	2SLS	2SLS	2SLS	2SLS
Dependent variable:	$ln\begin{pmatrix} Conflict \\ Events+1 \end{pmatrix}$	$ln\begin{pmatrix} Intrastate \\ Events+1 \end{pmatrix}$	$ln\begin{pmatrix} State\ vs.\ Rebels \\ Events+1 \end{pmatrix}$	$ln(Deaths+1)$
	(1)	(2)	(3)	(4)
Dams	0.055^{b}	0.054^{b}	0.040^{b}	0.069^{b}
	(0.023)	(0.023)	(0.018)	(0.033)
Basins	15923	15923	15923	15923
Observations	445844	445844	445844	445844
Kleibergen-Paap F-statistic	26.089	26.089	26.089	26.089
Mean dep. var.	.052	.051	.036	.091
Basin FE	✓	✓	✓	✓
Country-year FE	✓	✓	✓	✓
Geographic controls	✓	✓	✓	✓

Notes: 2SLS estimations. The unit of observation is a drainage basin and a year. The sample includes 15,923 basins for the years 1989-2016. The dependent variables measure conflict intensity, by taking the logarithm of the number of conflict events ln(Conflict Events+1). Column 1: "Events" is the sum of the reported conflict events in a basin and year. Column 2 and 3: the sum of events classified as "Intrastate" and "State vs. Rebels" are considered. Column 4: the sum of reported fatalities are considered. "Dams" specifies the stock of dams in a basin and year. Geographic controls: elevation, gradient, river length and basin size, interacted with $\overline{D}_{ct}$. The regressions control for basin and country-year fixed effects. Coefficients are reported with spatially clustered standard errors in parentheses, allowing for a spatial correlation within a 400 km radius of a basin's centroid and infinite serial correlation (Conley, 1999). ^{c} significant at 10%; ^{b} significant at 5%; ^{a} significant at 1%.

Table A5: Conflict Intensity: Small Events Versus Large Events

Estimation:	2SLS	2SLS	2SLS	2SLS
Dep. var. - Conflict intensity:	$(x \leq P_{25})$	$(P_{25} < x \leq P_{50})$	$(P_{50} < x \leq P_{75})$	$(x < P_{75})$
	(1)	(2)	(3)	(4)
Dams	0.012^{b}	0.012^{b}	0.008^{c}	0.000
	(0.006)	(0.005)	(0.004)	(0.004)
Basins	15923	15923	15923	15923
Observations	445844	445844	445844	445844
Kleibergen-Paap F-statistic	26.089	26.089	26.089	26.089
Mean dep. var.	.01	.008	.008	.008
Basin FE	✓	✓	✓	✓
Country-year FE	✓	✓	✓	✓
Geographic controls	✓	✓	✓	✓

Notes: 2SLS estimations. The unit of observation is a drainage basin and a year. The sample includes 15,923 basins for the years 1989-2016. The dependent variables indicate conflict incidence and are equal one if at least one conflict event occurs in a basin and year. In detail, conflict incidents are divided into four sub-categories by the intensity of battle-related deaths. Categories are separated at the 25th, 50th and 75th percentile, corresponding to 4, 13 and 50 fatalities. E.g. the dependent variable in column 4 equals one, if more than 50 casualties are recorded in a basin and year. "Dams" specifies the stock of dams in a basin and year. Geographic controls: elevation, gradient, river length and basin size, interacted with $\overline{D}_{ct}$. The regressions control for basin and country-year fixed effects. Coefficients are reported with spatially clustered standard errors in parentheses, allowing for a spatial correlation within a 400 km radius of a basin's centroid and infinite serial correlation (Conley, 1999). ^{c} significant at 10%; ^{b} significant at 5%; ^{a} significant at 1%.

Table A6: Controlling for Lagged Dependent Variable

Estimation:	2SLS	2SLS	2SLS	2SLS	2SLS
Dependent variable:	Conflict	Conflict	Conflict	Intrastate	State vs. Rebels
	(1)	(2)	(3)	(4)	(5)
Lagged dependent variable	0.259[a]	0.259[a]	0.259[a]	0.262[a]	0.279[a]
	(0.008)	(0.008)	(0.008)	(0.008)	(0.009)
Dams	-0.005[c]		0.021[b]	0.020[b]	0.018[b]
	(0.003)		(0.010)	(0.009)	(0.008)
$RG_i \times \overline{D}_{ct}$ (× 1000 *)		0.056[b]			
		(0.022)			
Basins	15923	15923	15923	15923	15923
Observations	429921	429921	429921	429921	429921
Kleibergen-Paap F-statistic	.	.	25.740	25.740	25.739
Mean dep. var.	.034	.034	.034	.034	.024
Basin FE	✓	✓	✓	✓	✓
Country-year FE	✓	✓	✓	✓	✓
Geographic controls	✓	✓	✓	✓	✓

Notes: The unit of observation is a drainage basin and a year. The sample includes 15,923 basins for the years 1990-2016. Column 1: OLS estimation. Column 2: reduced form estimation, with the coefficient and standard error multiplied by 1,000 (*). Columns 3-5: 2SLS estimations, with the first-stage Kleibergen-Paap F-statistics reported. The dependent variables indicate conflict incidence and are equal one if at least one conflict event occurs in a basin and year. "Conflict" considers all conflict events; "Intrastate" excludes events involving two state actors; "State vs. Rebels" only includes conflict events in which government forces fight against rebel groups. "Lagged dependent variable" controls for the conflict incident in period $t-1$. "Dams" specifies the stock of dams in a basin and year. Geographic controls: elevation, gradient, river length and basin size, interacted with $\overline{D}_{ct}$. The regressions control for basin and country-year fixed effects. Coefficients are reported with spatially clustered standard errors in parentheses, allowing for a spatial correlation within a 400 km radius of a basin's centroid and infinite serial correlation (Conley, 1999). [c] significant at 10%; [b] significant at 5%; [a] significant at 1%.

Table A7: The Timing of Violence

Estimation:	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS
Dependent variable:	Conflict	Conflict	Conflict	Intrastate	Intrastate	Intrastate
	(1)	(2)	(3)	(4)	(5)	(6)
Dams	0.050^b			0.048^b		
	(0.023)			(0.022)		
Dams (construction period excluded)		0.064^b			0.062^b	
		(0.029)			(0.028)	
Dams (construction + post completion)			0.057^b			0.055^b
			(0.028)			(0.027)
Basins	15544	15464	15544	15544	15464	15544
Observations	435232	432999	435232	435232	432999	435232
Kleibergen-Paap F-statistic	14.812	14.533	10.426	14.812	14.533	10.426
Mean dep. var.	.034	.034	.034	.034	.034	.034
Basin FE	✓	✓	✓	✓	✓	✓
Country-year FE	✓	✓	✓	✓	✓	✓
Geographic controls	✓	✓	✓	✓	✓	✓

Notes: 2SLS estimations. The unit of observation is a drainage basin and a year. The sample includes all untreated basins and treated basins for which the construction start year could be identified. The dependent variables indicate conflict incidence and are equal one if at least one conflict event occurs in a basin and year. "Conflict" considers all conflict events; "Intrastate" excludes events involving two state actors; "State vs. Rebels" only includes conflict events in which government forces fight against rebel groups. "Dams" specifies the stock of dams in a basin and year, where dams are considered from the year of completion onward. "Dams (construction period excluded)" excludes the years of dam construction from the sample. "Dams (construction + post completion)" reports a new in the data from the first construction year onward. Geographic controls: elevation, gradient, river length and basin size, interacted with $\overline{D}_{ct}$. The regressions control for basin and country-year fixed effects. Coefficients are reported with spatially clustered standard errors in parentheses, allowing for a spatial correlation within a 400 km radius of a basin's centroid and infinite serial correlation (Conley, 1999). ^{c} significant at 10%; ^{b} significant at 5%; ^{a} significant at 1%.

Table A8: Assessing the Role of Spatial Spillovers

Estimation:	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS
Dependent variable:	Conflict	Conflict	Conflict	Conflict	Conflict	Conflict
	(1)	(2)	(3)	(4)	(5)	(6)
Dams		0.041^b	0.050^b	0.046^b	0.029	0.049^b
		(0.018)	(0.020)	(0.022)	(0.069)	(0.022)
Dams in upstream basin	0.019		-0.011^b	-0.006	0.009	
	(0.044)		(0.005)	(0.010)	(0.066)	
Basins	12519	12519	12519	12519	12519	12473
Observations	350532	350532	350532	350532	350532	349244
Kleibergen-Paap F-statistic	6.332	17.417	16.121	7.053	1.291	14.905
Basin FE	✓	✓	✓	✓	✓	✓
Country-year FE	✓	✓	✓	✓	✓	✓
Geographic controls - own basin	✓	✓	✓	✓	✓	✓
Upstream dams instrumented	✓			✓	✓	
Geographic controls - upstream	✓				✓	

Notes: The unit of observation is a drainage basin and a year. The sample includes 12,517 basins for the years 1989-2016. The dependent variables indicate conflict incidence and are equal one if at least one conflict event occurs in a basin and year. "Conflict" considers all conflict events; "Intrastate" excludes events involving two state actors; "State vs. Rebels" only includes conflict events in which government forces fight against rebel groups. "Dams" specifies the stock of dams in a basin and year and is always instrumented for. "Dams in upstream basin" specifies the stock of dams in a basin's upstream basin and year and is instrumented for where indicated. Geographic controls of a basin (upstream basins): elevation, gradient, river length and basin size, interacted with $\overline{D}_{ct}$ ($\overline{D}^U_{ct}$). The regressions control for basin and country-year fixed effects. Coefficients are reported with spatially clustered standard errors in parentheses, allowing for a spatial correlation within a 400 km radius of a basin's centroid and infinite serial correlation (Conley, 1999). ^{c} significant at 10%; ^{b} significant at 5%; ^{a} significant at 1%.

Table A9: Extensive Margin and Pre-Sample Treatment

Estimation:	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS
Dependent variable:	Conflict	Intrastate	State vs. Rebels	Conflict	Intrastate	State vs. Rebels
	(1)	(2)	(3)	(4)	(5)	(6)
First dam	0.070^a	0.068^a	0.060^a			
	(0.026)	(0.026)	(0.022)			
Dams				0.038^b	0.037^b	0.032^b
				(0.017)	(0.016)	(0.014)
Basins	15832	15832	15832	13926	13926	13926
Observations	443303	443303	443303	389928	389928	389928
Kleibergen-Paap F-statistic	27.883	27.883	27.883	22.667	22.667	22.667
Mean dep. var.	.034	.034	.024	.034	.034	.024
Basin FE	✓	✓	✓	✓	✓	✓
Country-year FE	✓	✓	✓	✓	✓	✓
Geographic controls	✓	✓	✓	✓	✓	✓

Notes: 2SLS estimations. The unit of observation is a drainage basin and a year. Columns 1-3: Treated observations are only considered for the first dam completed during the sample period, indicated by "First dam" Columns 4-6: Basins with dams in the pre-sample are excluded from the sample and "Dams" specifies the stock of dams in the remaining basins. The dependent variables indicate conflict incidence and are equal one if at least one conflict event occurs in a basin and year. "Conflict" considers all conflict events; "Intrastate" excludes events involving two state actors; "State vs. Rebels" only includes conflict events in which government forces fight against rebel groups. Geographic controls: basin size, elevation, average basin gradient and river length, interacted with $\overline{D}_{ct}$. The regressions control for basin and country-year fixed effects. Coefficients are reported with spatially clustered standard errors in parentheses, allowing for a spatial correlation within a 400 km radius of a basin's centroid and infinite serial correlation (Conley, 1999). ^{c} significant at 10%; ^{b} significant at 5%; ^{a} significant at 1%.

Table A10: The Main Sub-types of Dams

Estimation:	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS
Dependent variable:	Conflict	Conflict	Conflict	Intrastate	Intrastate	Intrastate
	(1)	(2)	(3)	(4)	(5)	(6)
Hydropower dams	0.010			0.011		
	(0.007)			(0.007)		
Irrigation dams		0.041			0.045	
		(0.029)			(0.029)	
Water supply dams			0.091			0.099
			(0.068)			(0.068)
Basins	15923	15923	15923	15923	15923	15923
Observations	445844	445844	445844	445844	445844	445844
Kleibergen-Paap F-statistic	57.432	27.818	11.838	57.432	27.818	11.838
Mean dep. var.	.034	.034	.034	.034	.034	.034
Basin FE	✓	✓	✓	✓	✓	✓
Country-year FE	✓	✓	✓	✓	✓	✓
Geographic controls	✓	✓	✓	✓	✓	✓

Notes: 2SLS estimations. The unit of observation is a drainage basin and a year. The sample includes 15,923 basins for the years 1989-2016. The most common sub-categories of dams, by their primary and secondary use, are considered. Columns 1 and 4 study hydropower dams; columns 2 and 5 irrigation; and columns 3 and 6 water supply dams. The dependent variables indicate conflict incidence and are equal one if at least one conflict event occurs in a basin and year. "Conflict" considers all conflict events; "Intrastate" excludes events involving two state actors. "Dams" specifies the stock of dams in a basin and year. $RG_i \times \overline{D}_{ct}$ is the instrument, where RG_i is a basin's combined river fraction with a gradient of 1.5-3% or >6%; $\overline{D}_{ct}$ is the number dams in a country and year. Geographic controls: elevation, gradient, river length and basin size, interacted with $\overline{D}_{ct}$. The regressions control for basin and country-year fixed effects. Coefficients are reported with spatially clustered standard errors in parentheses, allowing for a spatial correlation within a 400 km radius of a basin's centroid and infinite serial correlation (Conley, 1999). ^{c} significant at 10%; ^{b} significant at 5%; ^{a} significant at 1%.

Table A11: Heterogenous Effects of Dams on Conflict: Country Characteristics

Estimation:	2SLS	2SLS	2SLS	2SLS	2SLS
Dependent variable:	Conflict	Conflict	Conflict	Conflict	Conflict
	(1)	(2)	(3)	(4)	(5)
Dams	0.030^{b}	0.029^{b}	0.041^{a}	-0.021	-0.019
	(0.015)	(0.013)	(0.015)	(0.019)	(0.027)
Dams × High democracy	0.017				
	(0.068)				
Dams × High Polarization (country)		0.056			
		(0.077)			
Dams × High income			-0.056		
			(0.039)		
Dams × High agricultural land share				0.066^{a}	
				(0.024)	
Dams × High agriculture to GDP share					0.063^{c}
					(0.033)
Basins	13543	15923	13365	13688	9624
Observations	379204	445844	374220	383264	269472
Kleibergen-Paap F-statistic	14.001	17.243	14.586	10.278	9.736
Mean dep. var.	.034	.034	.034	.037	.044
Mean dep. var. - interaction group	.031	.049	.021	.04	.055
Basin FE	✓	✓	✓	✓	✓
Country-year FE	✓	✓	✓	✓	✓
Geographic controls	✓	✓	✓	✓	✓

Notes: 2SLS estimations. The unit of observation is a drainage basin and a year. The table tests country heterogeneous effects. Countries included in the sample depend on data availability of the tested heterogeneity. "High" indicates the sub-sample of countries with an above-median value in 1988 (pre-sample) of the tested heterogeneity, which is interacted with the number of dams per basin and year. Column 1: Heterogeneity across the level of democratization, tested with data from the Polity 4 Project. Column 2: Heterogeneity across the level ethno-linguistic polarization, based on the polarization formula from Montalvo and Reynal-Querol (2005) and tested with data from the Ethnologue. Column 3: Heterogeneity across the level of real income per capita, tested with data from the Penn World Table. Column 4: Heterogeneity across the level of agricultural land share, tested with data from Faostat. Column 5: Heterogeneity across the level of GDP generated in the agricultural sector, tested with data from the World Development Indicators. The dependent variables indicate conflict incidence and are equal one if at least one conflict event occurs in a basin and year. "Conflict" considers all conflict events. "Dams" specifies the stock of dams in a basin and year. "Dams × Heterogeneity" indicates the interaction term between the number of dams and the country-wide heterogeneity. Geographic controls: basin size, elevation, average basin gradient and river length, interacted with $\overline{D}_{ct}$. The regressions control for basin and country-year fixed effects. Coefficients are reported with spatially clustered standard errors in parentheses, allowing for a spatial correlation within a 400 km radius of a basin's centroid and infinite serial correlation (Conley, 1999). ^{c} significant at 10%; ^{b} significant at 5%; ^{a} significant at 1%.

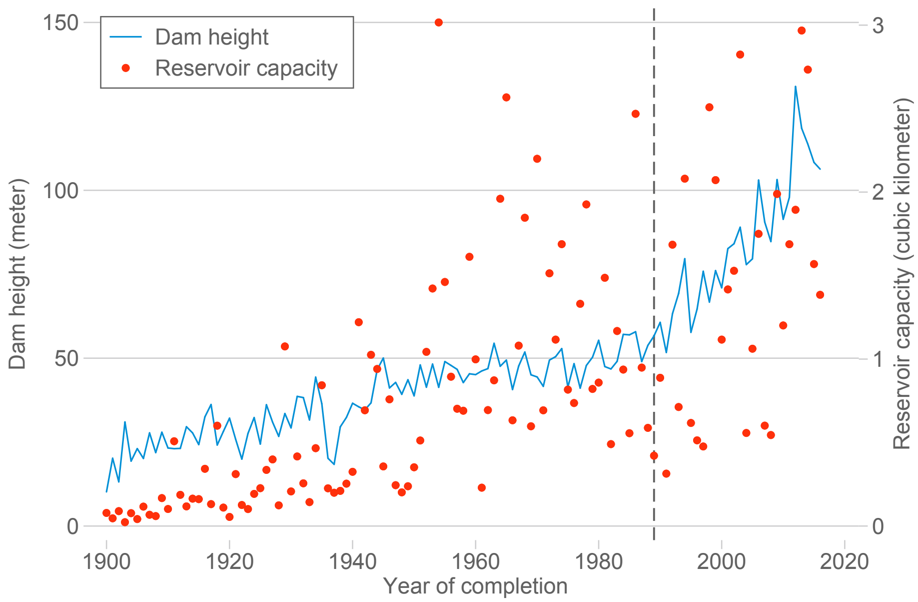
Table A12: Channel Analyis: Balancing Analysis (Cross-section)

Estimation:	2SLS	2SLS	2SLS	2SLS
Dependent variable	Low participation comp.	Low elect. comp.	High Polar. (basin)	High Fract. (basin)
	(1)	(2)	(3)	(4)
RG_i	-0.1679	-0.2939^b	-0.0077	-0.0266
	(0.1330)	(0.1348)	(0.0719)	(0.0730)
Basins	13543	13591	12225	12225
Observations	13543	13591	12225	12225
Mean dep. var	.628	.524	.458	.478
Country-year FE	✓	✓	✓	✓
Geographic controls	✓	✓	✓	✓

Notes: 2SLS estimations. The unit of observation is a drainage basin. The dependent variables indicate countries with low political competition in 1988 (columns 1-2) and high within-country polarization and fractionalization in 1975 (columns 3-4). Geographic controls: basin size, elevation, average basin gradient and river length. The regressions in columns 3-4 control for country fixed effects. Coefficients are reported with spatially clustered standard errors in parentheses, allowing for a spatial correlation within a 400 km radius of a basin's centroid. ^{c} significant at 10%; ^{b} significant at 5%; ^{a} significant at 1%.

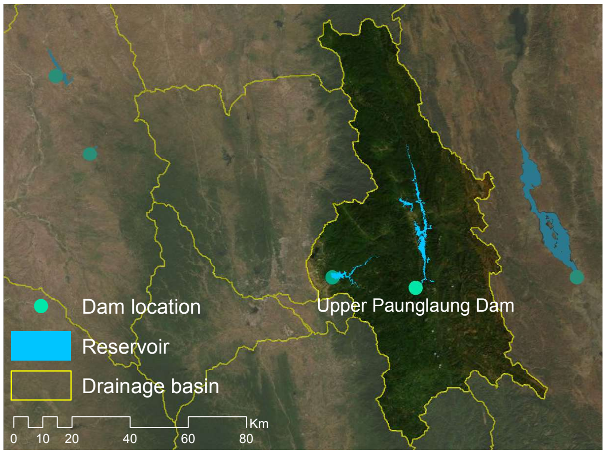
B Figures

Figure A1: Average Size of Newly-Constructed Dams



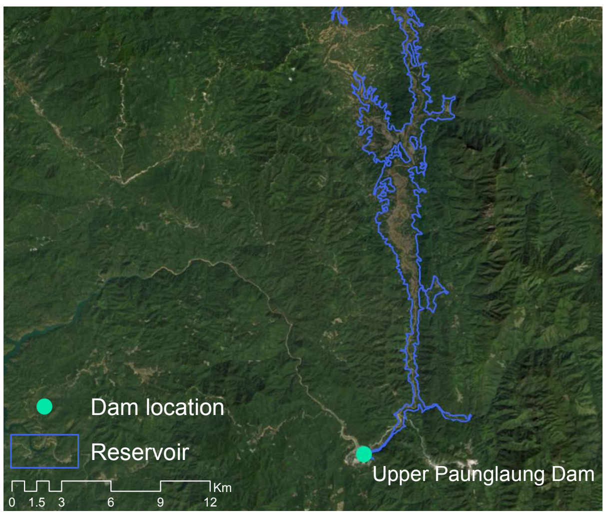
Notes: The graph is based on data from the Global Reservoir and Dam (GRanD) Database, depicting all "large" dams exceeding a reservoir capacity of 0.1 cubic kilometers. The blue solid line indicates the annual average dam height in meter of newly constructed dams (left scale); the red dots indicate the average annual reservoir capacity in cubic kilometer of newly constructed dams (right scale). The vertical black dashed line further indicates the first year of the sample (1989).

Figure A2: Data example: Upper Paunlaung Dam, Myanmar completed in 2015



A) Drainage basin of Paunlaung river section.

B) Data precision, satellite image from 2018.

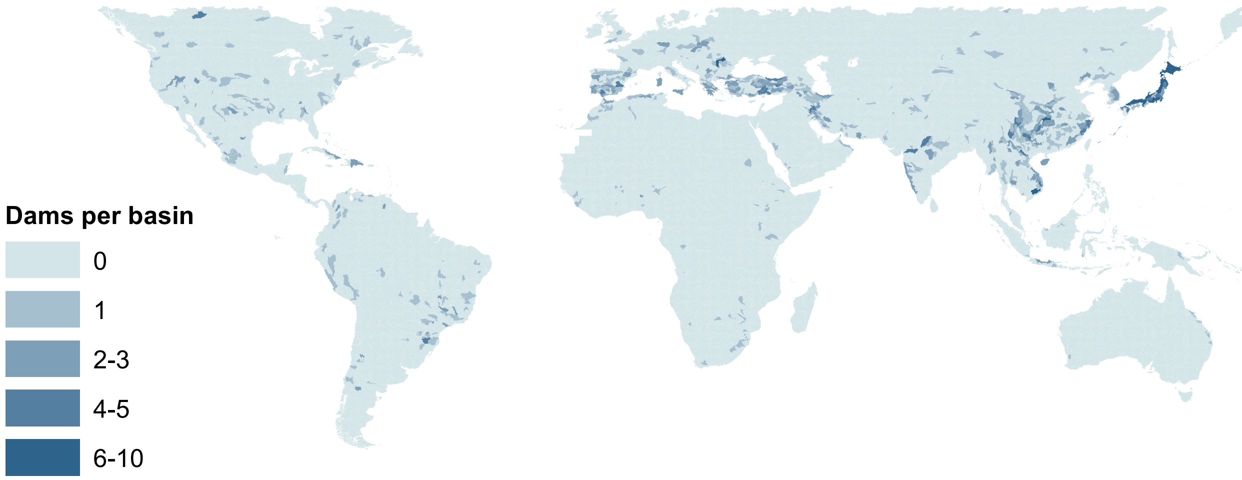


C) Planned reservoir (blue) and pre-completion satellite image.

D) Planned reservoir (shaded) and pre-completion satellite image, magnified.

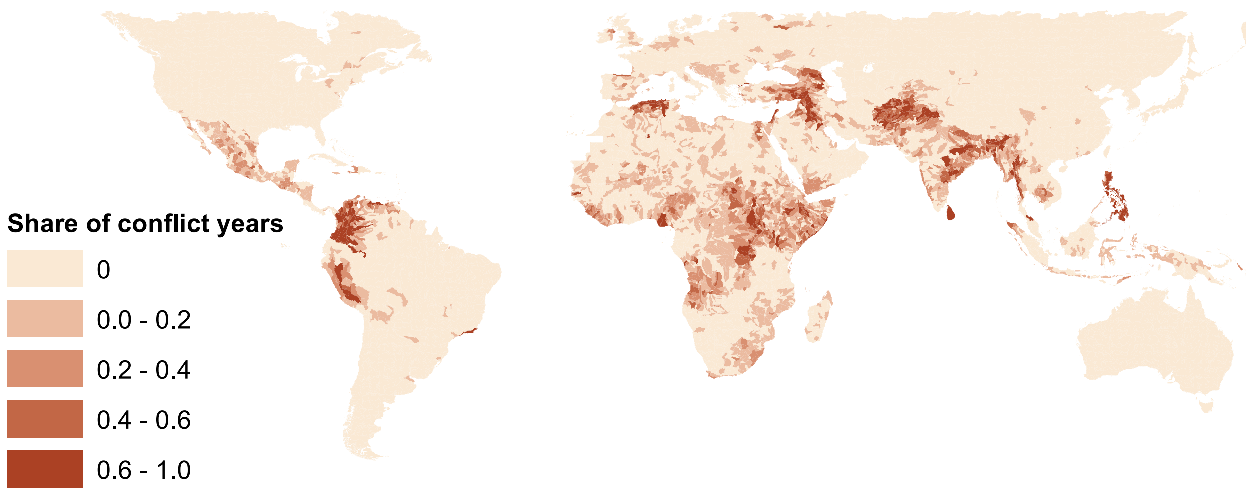
Notes: The satellite images depict Upper Paunlang Dam in Myanmar, prior and past dam construction. Data on dam location and reservoir extent is based on data from GranD. The image was produced ArcGIS (ESRI).

Figure A3: Spatial Distribution of Dams, 1989-2016



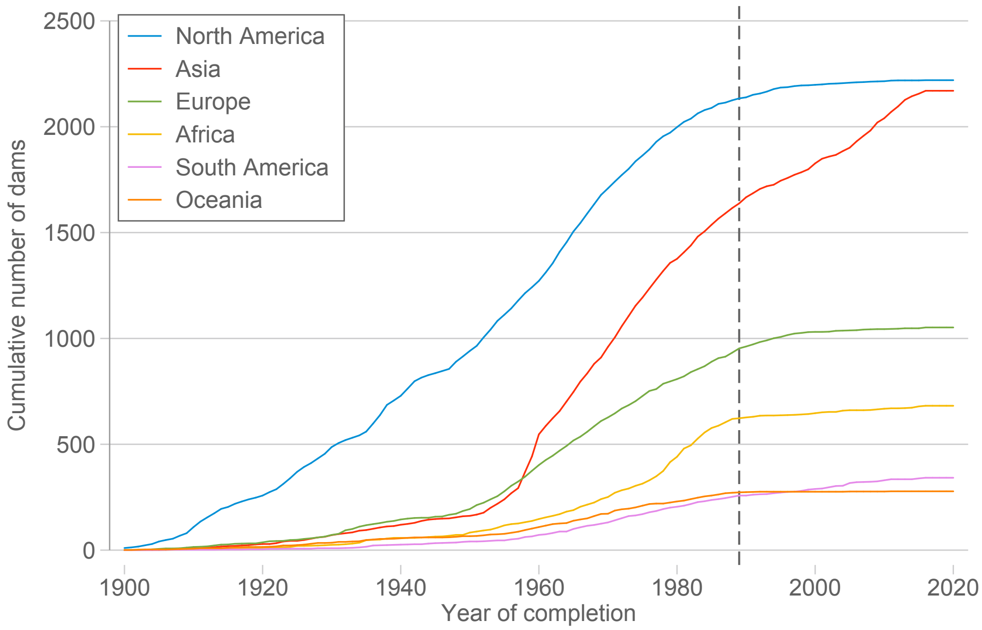
Notes: This figure depicts the spatial distribution large dams build between 1989-2016, with data from GRanD. Darker shadings indicate basins with a higher number of dams.

Figure A4: Spatial Distribution of Conflict, 1989-2016



Notes: This figure depicts the spatial distribution conflict incidents between 1989-2016, with data from UCDP - GED. Darker shadings indicate basins with a higher share of violent sample years.

Figure A5: Dams by Continent



Notes: Dams are displayed by continent since 1900 with data from GRanD. The vertical dashed black line indicates the start of the sample, 1989.

C Robustness of the instrument

This section addresses potential concerns about the instrument with a battery of robustness exercises.

C.1 Revisiting $\overline{D}_{ct}$

Table A13: Excluding the Smallest 10% of Countries / Basins

Estimation:	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS
Dependent variable:	Conflict	Intrastate	Conflict	Intrastate	Conflict	Intrastate
	(1)	(2)	(3)	(4)	(5)	(6)
Dams	0.032^b	0.031^b	0.028^b	0.027^b	0.028^b	0.027^b
	(0.013)	(0.013)	(0.012)	(0.011)	(0.013)	(0.012)
Basins	15862	15862	14279	14279	12692	12692
Observations	444136	444136	399812	399812	355376	355376
Kleibergen-Paap F-statistic	26.088	26.088	29.649	29.649	33.308	33.308
Mean dep. var.	.034	.034	.034	.034	.034	.034
Basin FE	✓	✓	✓	✓	✓	✓
Country-year FE	✓	✓	✓	✓	✓	✓
Geographic controls	✓	✓	✓	✓	✓	✓
Exclude smallest 10% of countries	✓	✓	✓	✓	✓	✓
Exclude smallest 10% of basins			✓	✓	✓	✓
Exclude largest 10% of basins					✓	✓

Notes: 2SLS estimations. The unit of observation is a drainage basin and a year. Columns 1-2 exclude 10% of countries with the least number of drainage basins. Columns 3-4 exclude the smallest 10% of drainage basins by surface and columns 5-6 exclude the largest 10% of drainage basins by surface. The dependent variables indicate conflict incidence and are equal one if at least one conflict event occurs in a basin and year. "Conflict" considers all conflict events; "Intrastate" excludes events involving two state actors; "State vs. Rebels" only includes conflict events in which government forces fight against rebel groups. "Dams" specifies the stock of dams in a basin and year. Geographic controls: basin size, elevation, average basin gradient and river length, interacted with $\overline{D}_{ct}$. The regression controls for basin and country-year fixed effects. Coefficients are reported with spatially clustered standard errors in parentheses, allowing for a spatial correlation within a 400 km radius of a basin's centroid and infinite serial correlation (Conley, 1999). ^{c} significant at 10%; ^{b} significant at 5%; ^{a} significant at 1%.

First, the time-variant part of the instrument $\overline{D}_{ct}$ measuring the stock of dams per country and year could suffer from endogeneity if the country-wide demand for dams is driven by a single drainage basin. This issue has been pointed out in previous work and turns out to be less problematic in the context of this paper. Unlike in existing studies, the "shift" part of my instrument does not vary at the subnational, but at the national level. Therefore, variation in $\overline{D}_{ct}$ stems from all regions of a country, reducing the risk of a single basin driving the total demand for dams in a country. Smaller countries with lower numbers of drainage basins could still cause such bias and I address this potential threat in columns 1 to 2 of Table A13, by excluding the ten percent of countries with the least number of drainage basins, i.e. with less than seven drainage basins per country. Results are widely unaffected, suggesting the estimation coefficients are unlikely to be biased by small countries. Results neither change when additionally excluding the smallest or largest ten percent of drainage basins by surface area, as shown in columns 3 to 4, and 5 to 6, respectively.

Table A14: Deriving D_{ct} from Past Dam Stocks (Moving Average)

Estimation:	2SLS					
$\overline{D}_{ct}$ bssed on:	5 year moving average			10 year moving average		
Dependent variable:	Conflict	Intrastate	State vs. Rebels	Conflict	Intrastate	State vs. Rebels
	(1)	(2)	(3)	(4)	(5)	(6)
Dams	0.044^a	0.043^a	0.037^a	0.051^b	0.050^b	0.051^a
	(0.016)	(0.016)	(0.013)	(0.021)	(0.021)	(0.017)
Basins	15923	15923	15923	15923	15923	15923
Observations	445844	445844	445844	445844	445844	445844
Kleibergen-Paap F-statistic	25.698	25.698	25.698	26.368	26.368	26.368
Mean dep. var.	.034	.034	.024	.034	.034	.024
Basin FE	✓	✓	✓	✓	✓	✓
Country-year FE	✓	✓	✓	✓	✓	✓
Geographic controls	✓	✓	✓	✓	✓	✓

Notes: 2SLS estimations. The unit of observation is a drainage basin and a year. The sample includes 15,923 basins for the years 1989-2016. The dependent variables indicate conflict incidence and are equal one if at least one conflict event occurs in a basin and year. "Conflict" considers all conflict events; "Intrastate" excludes events involving two state actors; "State vs. Rebels" only includes conflict events in which government forces fight against rebel groups. "Dams" specifies the stock of dams in a basin and year. The time-variant part of the instrument of the stock of dams $\overline{D}_{ct}$ is replaced with the moving average of a country's dam stock over the past five and ten periods, in columns 1-3 and 4-6, respectively. Geographic controls: basin size, elevation, average basin gradient and river length, interacted with the moving average of $\overline{D}_{ct}$. The regressions control for basin and country-year fixed effects. Coefficients are reported with spatially clustered standard errors in parentheses, allowing for a spatial correlation within a 400 km radius of a basin's centroid and infinite serial correlation (Conley, 1999). ^{c} significant at 10%; ^{b} significant at 5%; ^{a} significant at 1%.

Another potential concern is that counting new dams towards $\overline{D}_{ct}$ from the year of dam completion onward may over-fit the econometric model. One way to address this issue is by deriving $\overline{D}_{ct}$ solely from the past stock of dams, hence excluding new dams from $\overline{D}_{ct}$.[37] Predicting current dams investment from the past appears conceptually valid, because a country's know-how in dam construction may reduce the costs of dams in the future. Columns 1 to 3 and 4 to 6 of Table A14 respectively replace $\overline{D}_{ct}$ with the five and ten year moving average of countries' past dams stocks. The results are in line with the baseline.

[37] Alternatively, one could base $\overline{D}_{ct}$ on the number of dams of neighboring states (assuming that states in proximity follow similar economic trajectories). Such approach does however not yield a sufficiently strong first stage.

Table A15: Cross-sectional Analysis

Estimation:	OLS	Reduced F.	2SLS	2SLS	2SLS
Dependent variable:	Conflict share	Conflict share	Conflict share	Intrastate	State vs. Rebels
	(1)	(2)	(3)	(4)	(5)
# Dams	0.0041		0.2401^b	0.2457^b	0.1920^b
	(0.0066)		(0.1147)	(0.1157)	(0.0978)
RG_i		0.0283^b			
		(0.0120)			
Basins	15923	15923	15923	15923	15923
Observations	15923	15923	15923	15923	15923
Kleibergen-Paap F-statistic	.	.	10.618	10.618	10.618
Mean dep. var.	.034	.034	.034	.034	.034
Geographic controls	✓	✓	✓	✓	✓
Country FE	✓	✓	✓	✓	✓

Notes: The unit of observation is a drainage basin. The sample includes 15,923 basins. Column 1: OLS estimation. Column 2: reduced form estimation, with the coefficient and standard error multiplied by 1,000 (*). Columns 3-5: 2SLS estimations, with the first-stage Kleibergen-Paap F-statistics reported. The dependent variables indicate the share of years between 1989 to 2016 with at least one conflict incidence. “Conflict share” considers all conflict events; “Intrastate share” excludes events involving two state actors; “State vs. Rebels share” only includes conflict events in which government forces fight against rebel groups. “# Dams” specifies the number of dams built per basin between 1989 to 2016. Geographic controls: basin size, elevation, average basin gradient and river length. The regression controls for country fixed effects. Coefficients are reported with spatially clustered standard errors in parentheses, allowing for a spatial correlation within a 400 km radius of a basin’s centroid. ^{c} significant at 10%; ^{b} significant at 5%; ^{a} significant at 1%.

To further address whether issues stem from the time-variant part of the instrument, a cross-sectional analysis is performed which ignores the time dimension of the analysis, and by this, omits $\overline{D}_{ct}$ from the econometric specification. In detail, I instrument the endogenous variable measuring the total of dams built per basin during 1989 to 2016 with the time-invariant river gradient part of the instrument, RG_i. A dependent variable is constructed that accounts for the share of periods between 1989 to 2016 with at least one conflict incident per basin, a standard definition in the conflict literature. The regressions further account for country fixed effects. Results are reported in Table A15 with positive and weakly significant coefficients for the 2SLS. While these findings are in line with the panel baseline discussed earlier, a cross-sectional investigation is less robust compared to the panel, due to the absence of basin fixed effects which potentially introduces omitted variable bias. Nevertheless, two important observations can be made: First, the cross-sectional analysis finds a similar association as the panel. Second, these results do not depend on the construction of $\overline{D}_{ct}$.

Table A16: Predicting D_{ct} from the Pre-Sample Continent Share

Estimation:	OLS	Reduced F.	2SLS	2SLS	2SLS
Dependent variable:	Conflict	Conflict	Conflict	Intrastate	State vs. Rebels
	(1)	(2)	(3)	(4)	(5)
Dams	-0.007^{b}		0.050^{b}	0.049^{b}	0.044^{b}
	(0.004)		(0.023)	(0.022)	(0.018)
$RG_i \times D_{ct}^{predicted}$ (× 1000 *)		0.183^{b}			
		(0.073)			
Basins	15923	15923	15923	15923	15923
Observations	445844	445844	445844	445844	445844
Kleibergen-Paap F-statistic	.	.	23.168	23.168	23.168
Mean dep. var.	.034	.034	.034	.034	.034
Basin FE	✓	✓	✓	✓	✓
Country-year FE	✓	✓	✓	✓	✓
Geographic controls	✓	✓	✓	✓	✓

Notes: The unit of observation is a drainage basin and a year. The sample includes 15,923 basins for the years 1989-2016. Column 1: OLS estimation. Column 2: reduced form estimation, with the coefficient and standard error multiplied by 1,000 (*). Columns 3-5: 2SLS estimations, with the first-stage Kleibergen-Paap F-statistics reported. The dependent variables indicate conflict incidence and are equal one if at least one conflict event occurs in a basin and year. "Conflict" considers all conflict events; "Intrastate" excludes events involving two state actors; "State vs. Rebels" only includes conflict events in which government forces fight against rebel groups. "Dams" specifies the stock of dams in a basin and year. $RG_i \times D_{ct}^{predicted}$ is the instrument, where RG_i is a basin's river fraction with a gradient of 1.5-3% or >6%; $D_{ct}^{predicted}$ is the predicted number dams in a country and year defined as a country's pre-sample continent share of dams multiplied by the actual number of dams per continent each sample year. Geographic controls: basin size, elevation, average basin gradient and river length, interacted with $D_{ct}^{predicted}$. The regressions control for basin and country-year fixed effects. Coefficients are reported with spatially clustered standard errors in parentheses, allowing for a spatial correlation within a 400 km radius of a basin's centroid and infinite serial correlation (Conley, 1999). ^{c} significant at 10%; ^{b} significant at 5%; ^{a} significant at 1%.

Next, I return to the panel and follow Duflo and Pande (2007) who predict the number of dams per country and year $D_{ct}^{predicted}$ from a country's pre-sample continent share of dams multiplied by the actual number of dams per continent each sample year. This gives an alternative instrument $RG_i \times D_{ct}^{predicted}$, with the river gradient multiplied with the predicted number of dams per country. Constructing the instrument in such fashion rules out a potential bias caused by all dams of a country located a single basin, as well as an over-fitting of the model. The 2SLS results in Table A16 show positive coefficients throughout, larger in magnitude than the baseline and significant at the five-percent level. The Kleibergen-Paap F-Statistic in column 3 reduces to 23.17, since using pre-sample country-variation introduces some noise. It is important to note that 6 out of 77 treated countries had no dams in the pre-sample period, which implies a $D_{ct}^{predicted}$ of zero. In other words, those countries do not comply with this instrument specification, which reduces the external validity of the findings. For that reason and the fact that a bias caused by selection of dams to a single basin appears unlikely, the remainder of the paper employs the actual number of dams $\overline{D}_{ct}$, as described in the baseline specification.

Table A17: Controlling for Basin-Specific Time Trends

Estimation:	OLS	Reduced F.	2SLS	2SLS	2SLS
Dependent variable:	Conflict	Conflict	Conflict	Intrastate	State vs. Rebels
	(1)	(2)	(3)	(4)	(5)
Dams	-0.007^c		0.031^a	0.030^a	0.027^a
	(0.003)		(0.011)	(0.011)	(0.009)
$RG_i \times \overline{D}_{ct}$ ($\times$ 1000 *)		0.084^a			
		(0.025)			
Basins	15923	15923	15923	15923	15923
Observations	445844	445844	445844	445844	445844
Kleibergen-Paap F-statistic	.	.	30.715	30.715	30.715
Mean dep. var.	.034	.034	.034	.034	.024
Basin-specific time trends	✓	✓	✓	✓	✓
Country-year FE	✓	✓	✓	✓	✓
Geographic controls	✓	✓	✓	✓	✓

Notes: The unit of observation is a drainage basin and a year. The sample includes 15,923 basins for the years 1989-2016. Column 1: OLS estimation. Column 2: reduced form estimation, with the coefficient and standard error multiplied by 1,000 (*). Columns 3-5: 2SLS estimations, with the first-stage Kleibergen-Paap F-statistics reported. The dependent variables indicate conflict incidence and are equal one if at least one conflict event occurs in a basin and year. “Conflict” considers all conflict events; “Intrastate” excludes events involving two state actors; “State vs. Rebels” only includes conflict events in which government forces fight against rebel groups. “Dams” specifies the stock of dams in a basin and year. $RG_i \times \overline{D}_{ct}$ is the instrument, where RG_i is a basin’s combined river fraction with a gradient of 1.5-3% or $> 6\%$; $\overline{D}_{ct}$ is the number dams in a country and year. Geographic controls: basin size, elevation, average basin gradient and river length, interacted with $\overline{D}_{ct}$. The regressions control for basin-specific time trends and country-year fixed effects. Coefficients are reported with standard errors clustered at the basin level in parentheses due to computational limitations. ^{c} significant at 10%; ^{b} significant at 5%; ^{a} significant at 1%.

The next exercise aims at addressing endogeneity stemming from differences in the basin-induced growth over time, by replacing basin fixed effects with linear basin-specific time trends. Results in Table A17 remain comparable to the baseline results. I believe the econometric specification of the baseline accounting for basin fixed effects superior, because of the difference-in-difference character that allows a straightforward interpretation of the coefficients.

C.2 Revisiting RG_i

Table A18: Alternative RG_i and Placebo Tests

Estimation:	2SLS	2SLS	2SLS	2SLS	OLS	OLS
Dependent variable:	Conflict	Conflict	Conflict	Conflict	Conflict share	Conflict share
	(1)	(2)	(3)	(4)	(5)	(6)
Dams	0.050^b	0.032^b	0.032^b	0.035^b		
	(0.020)	(0.013)	(0.013)	(0.014)		
RG_i					0.033^b	-0.010
					(0.013)	(0.019)
Basins	15923	15923	15923	15923	12167	3756
Observations	445844	445844	445844	429921	12167	3756
Kleibergen-Paap F-statistic	17.117	25.180	25.480	26.471	.	.
Mean dep. var.	.034	.034	.034	.034	.034	.034
Data	Panel	Panel	Panel	Panel	Cross-section	Cross-section
Geographic controls	✓	✓	✓	✓	✓	✓
Basin FE	✓	✓	✓	✓		
Country-year FE	✓	✓	✓	✓		
Gradient in IV band control		✓	✓	✓		
Large rivers control			✓	✓		
ln(Population+1) control				✓		
Country FE					✓	✓

Notes: Columns 1-4: 2SLS estimations. The unit of observation is a drainage basin and a year. The dependent variable indicates conflict incidence and is equal one if at least one conflict event occurs in a basin and year. Geographic controls: basin size, elevation, average basin gradient and river length, interacted with $\overline{D}_{ct}$. The regressions control for basin and country-year fixed effects. "Dams" specifies the stock of dams in a basin and year. Columns 1: RG_i is defined as a basin's area share (instead of the gradient just along the river center line) with a gradient of 1.5-3% and >6%. Column 2: As baseline, additionally controls for the average gradient along rivers. Column 3: As column 2, additionally controlling for large rivers with data from Natural Earth. Column 4: As column 4, additionally controlling for the log of the basin population plus 1, with data from the Global Human Settlement Layer (for missing years, data is linearly interpolated). Coefficients are reported with spatially clustered standard errors in parentheses, allowing for a spatial correlation within a 400 km radius of a basin's centroid and infinite serial correlation (Conley, 1999). Columns 4-5: OLS estimations. The unit of observation is a drainage basin. The dependent variable is the share of years with conflict incidents. Placebo tests performing the reduced form estimations on sub-samples. Column 5: Sub-sample of countries with at least one new dam (77 countries). Column 6: Sub-sample of countries without new dams (76 countries). Geographic controls: basin size, elevation, average basin gradient and river length. The regressions control for country fixed effects. Coefficients are reported with spatially clustered standard errors in parentheses, allowing for a spatial correlation within a 400 km radius of a basin's centroid and, in the panel setting of columns 1-4, infinite serial correlation (Conley, 1999). ^{c} significant at 10%; ^{b} significant at 5%; ^{a} significant at 1%.

In what follows, alternative versions of the time-invariant part of the instrument RG_i are considered. In the baseline specification, RG_i is chosen to capture the gradient just along the river line. Defining a wider band could run the risk to violate the exclusion restriction by introducing faraway gradient characteristics irrelevant for dams, but not for conflict. Nevertheless, previous papers have defined the river gradient across complete spatial units (e.g. basin or districts). In column 1 of Table A18, I follow this procedure and define RG_i across the whole basin surface. The pairwise correlation with the baseline river gradient is 0.98 and the correlation coefficient statistically significant at the 1 percent level. The coefficient of 0.05 exceeds the baseline and is significant at the 5 percent level. Further, a reduction in Kleibergen-Paap F-Statistics might be due to additional noise introduced via geographic features irrelevant for predicting dam placement.

Returning to the baseline specification, a set of additional robustness exercises is performed. Column 2 of Table A18 controls for the average gradient along a basin's river lines, to capture potential confounding factors on the same geographic extent on which RG_i is based on. And to address a potential bias from differently-sized rivers, column 3 further controls for length of large rivers in a basin, with data on the world's 1000 largest from Natural Earth (National Geographic, 2019). The additional controls in columns 2 and 3 are interacted with $\overline{D}_{ct}$, and the reported results remain widely unchanged. Next, I additionally control for the basin population, with data from the Global Human Settlement Layer available for the years 1975, 1990, 2000 and 2015 (Schiavina, Freire and Mac Manus, 2019). Data is linearly interpolated for missing years, and the year 2016 is coded as missing. The robustness exercises on RG_i conclude with a set of placebo tests to demonstrate that the instrument is unrelated to conflict in places where dams have not been built. In detail, I split the sample into countries with at least one new dam (77 countries) and untreated countries (76 countries) and run for each sub-sample reduced form regressions in columns 5 and 6, respectively. Untreated basins have no variation in the number of dams, I therefore move to the cross-section and regress a basin's share of sample years with conflict on RG_i. While the river gradient shows a five percent significant and positive association with the conflict share in treated countries, the effect in basins of untreated countries is statistically insignificant. It is reassuring to yield results in line with the expected, having in mind that the fixed effect structure in this cross-sectional setting only accounts for country fixed effects and is hence less restrictive than the baseline.[38]

[38]The sample is split by treated countries rather than treated basins, because spatial spillovers within

Table A19: Three Excluded Instruments

Estimation:	OLS	Reduced F.	2SLS	2SLS	2SLS
Dependent variable:	Conflict	Conflict	Conflict	Intrastate	State vs. Rebels
	(1)	(2)	(3)	(4)	(5)
Dams	-0.007^{c}		0.024^{b}	0.023^{b}	0.022^{b}
	(0.004)		(0.012)	(0.011)	(0.010)
Gradient $1.5-3\%$ $\times\overline{D}_{ct}$ ($\times$ 1000 *)		0.102^{a}			
		(0.038)			
Gradient $3-6\%$ $\times\overline{D}_{ct}$ ($\times$ 1000 *)		0.009			
		(0.037)			
Gradient $>6\%$ $\times\overline{D}_{ct}$ ($\times$ 1000 *)		0.053			
		(0.043)			
Basins	15923	15923	15923	15923	15923
Observations	445844	445844	445844	445844	445844
Kleibergen-Paap F-statistic	.	.	12.208	12.208	12.208
Mean dep. var.	.034	.034	.034	.034	.024
Basin FE	✓	✓	✓	✓	✓
Country-year FE	✓	✓	✓	✓	✓
Geographic controls	✓	✓	✓	✓	✓

Notes: The unit of observation is a drainage basin and a year. The sample includes 15,923 basins for the years 1989-2016. Column 1: OLS estimation. Column 2: reduced form estimation, with the coefficient and standard error multiplied by 1,000 (*). Columns 3-5: 2SLS estimations, with the first-stage Kleibergen-Paap F-statistics reported. The dependent variables indicate conflict incidence and are equal one if at least one conflict event occurs in a basin and year. "Conflict" considers all conflict events; "Intrastate" excludes events involving two state actors; "State vs. Rebels" only includes conflict events in which government forces fight against rebel groups. "Dams" specifies the stock of dams in a basin and year. The first-stage includes three instruments, one for each river gradient category's surface share, each interacted with $\overline{D}_{ct}$ is the number dams in a country and year. Geographic controls: basin size, elevation, average basin gradient and river length, interacted with $\overline{D}_{ct}$. The regressions control for basin and country-year fixed effects. Coefficients are reported with spatially clustered standard errors in parentheses, allowing for a spatial correlation within a 400 km radius of a basin's centroid and infinite serial correlation (Conley, 1999). ^{c} significant at 10%; ^{b} significant at 5%; ^{a} significant at 1%.

Next, the baseline is performed with three excluded instruments, one accounting for each of the river gradient classes discussed earlier. Results are depicted in Table A19 with a Kleibergen-Paap F-statistic of 12.21 and positive coefficients, just short of the baseline and significant at the five percent level. One shortcoming of this approach is that it does not necessarily capture the dual requirement of moderate and at the same time steep gradients

countries cannot fully be ruled out (c.f. Section 6.4), whereas a spread of violence across country borders is rather unlikely (c.f. Section 6.1.1). Further, the regression controls for un-interacted geographic controls.

for multi-purpose dams.[39]

[39] In addition, the just-identified baseline setting can be estimated efficiently with OLS, which allows to make use of the acreg Stata command to spatially clusters standard errors in the instrumental variable specification (Colella et al., 2019)

D The local economic impact of dams

Table A20: Local Economic Impact of Dams

Estimation:	2SLS	2SLS	2SLS	2SLS
Dependent variable:	ln(GDP/area)	ln(light/area)	ln(max{light/area,ϵ})	ln(Agriculture/area)
	(1)	(2)	(3)	(4)
Dams	-0.963	-0.675	-0.148	0.194
	(0.623)	(0.808)	(1.388)	(0.177)
Basins	2076	8468	12173	13999
Observations	58148	237114	340868	391987
Mean dep. var.	3.458	-.808	-3.813	-1.819
K.-P. F-stat.	19.932	18.906	26.813	24.930
Basin FE	✓	✓	✓	✓
Country-year FE	✓	✓	✓	✓
Geographic controls	✓	✓	✓	✓

Notes: 2SLS estimations. The unit of observation is a drainage basin and a year. The sample sizes vary with the data availability of the dependent variable. Column 1: The dependent variable measures the income per km^2 with "Gross Cell Products" data from G-Econ, available for 1990, 1995, 2000 and 2005. Column 2-3: Dependent variables measure the night light density per km^2 with data from NOAA, available for 1992-2013. Column 3 replaces zero-observations with $\epsilon = 0.000023$ (the smallest positive value in the sample.). Column 4: Dependent variable measures agricultural productivity per km^2 via the Normalized Difference Vegetation Index with data from NOAA. To proxy agricultural productivity, the data is limited regions with agricultural land. "Dams" specifies the stock of dams in a basin and year. Geographic controls: basin size, elevation, average basin gradient and river length, interacted with $\overline{D}_{ct}$. The regressions control for basin and country-year fixed effects. Coefficients are reported with spatially clustered standard errors in parentheses, allowing for a spatial correlation within a 400 km radius of a basin's centroid and infinite serial correlation (Conley, 1999). ^{c} significant at 10%; ^{b} significant at 5%; ^{a} significant at 1%.

To empirically assess the effect of dams on the local economy across the world, I rely on two main data sources. The first source is the "Gross Cell Product" from the G-Econ database, a dataset at the 1×1 decimal degree spatial resolution, or about 110×110 km at the equator (Nordhaus, 2006). This data is derived from regional income and employment registers and available for the years 1990, 1995, 2000 and 2005. Column 1 of Table A20 performs the 2SLS specification with the logarithm of the mean income per square kilometer as the dependent variable. The negative, although insignificant coefficient depicted in column 1 may hint towards a sizable, negative economic impact of dams in their direct surroundings, compared to counterfactual observations. The magnitude of the coefficient might be surprising at first, but a closer look reveals that frequently a sizable part of the local economy clustered along a river is submerged by the reservoir. For example, when the Upper Paunlaung dam in

Myanmar started operations in 2015, the reservoir eradicated most of the local economy, including 23 villages centered along the Paunlaung river banks and displaced around 10,000 people, as shown in Panel C and D of Figure A2 in the Appendix. Overall, the spatial resolution and frequency of the appear far from ideal to measure fine-grained year-to-year changes in the local economy.

The second data source, satellite night lights from NOAA (National Oceanic and Atmospheric Administration, 2013), is available for the years 1992 to 2013 with a comparatively high spatial resolution of 1×1 km. Night time satellite data have been used to approximate economic activity and electrification. In detail, the logarithm of the average light density per area has been shown to correlate linearily with the logarithm of total GDP (Henderson, Storeygard and Weil, 2012).[40] I calculate the logarithm of the average light density of each basin and year and results are shown in column 2 of Table A20. The coefficient measuring the effect of dams on local light density is also negative and insignificant.[41] The existing literature employing satellite nighttime lights commonly replaces observations with a zero light density by some small, positive number, because regions with a low - but non-zero - economic activity often have a reported zero light density, an issue known as *bottom censoring.* I follow Henderson et al. (2017) and replace all observations with zero light density with the smallest, positive amount in the sample, $\epsilon = 0.000023$. The coefficient reported in column 3 remains negative and insignificant.[42] Having in mind the limitations of the both data, the findings indicate that dams may tend to negatively impact their surrounding economies.

The last part of this section focuses on the effect of dams on agricultural productivity. Agriculture is a key economic sector in dammed basins, with an average of 41 percent of the basins covered with agricultural land, compared to 25 percent in the rest of the sample (Latham et al., 2014). The major challenge in quantitatively assessing the impact of dams on land productivity is the absence of globally consistent land survey panel data. Instead, I rely on Normalized Difference Vegetation Index (NDVI), a remotely-sensed data measuring

[40]Recent papers using these data include Henderson, Storeygard and Weil (2012); Hodler and Raschky (2014); Michalopoulos and Papaioannou (2014) and (Storeygard, 2016).

[41]To capture the effect of dams on the traditional economy, a 1 km buffer around dams is omitted from the analysis, because large dam construction sites may emit enough light be visible on satellite images.

[42]It turns out that the magnitude of the coefficient is sensitive to the choice of ϵ, suggesting that bottom censoring may be a serious concern when measuring the economic impact of dams via night lights in less developed areas.

local vegetational strength (Vermote et al., 2014). To approximately measure the evolution of agricultural land productivity, I restrict the data to regions classified as agricultural land as of 2014, i.e. close towards the end of the sample. Column 4 of Table A20 shows a statistically insignificant increase in vegetational productivity in the remaining farm land after dam completion.

A possible increase in agricultural productivity is plausible, because the presence of more water from the reservoir and from irrigation facilities ensures steady water supplies and allows for the cultivation of water intensive crops. A higher agricultural productivity associated with dams has been argued to distort local crop markets and potentially contribute to local income inequality, because it allows farmers to switch from food to exportable cash crops (Duflo and Pande, 2007). Contrarily, it could be possible that an increase in agricultural productivity mitigates the loss of farm land caused by the reservoir. To determine which of the two cases is true, more detailed poverty data would be required, which is however unavailable at the global scale. To summarize, these results appear to be in line with previous work, arguing that communities directly affected by dams are more likely to suffer economic losses. Conceptually speaking, negative income shocks could lower opportunity costs to join rebel forces, hence provide a plausible explanation for a surge in violence.

References

Abadie, Alberto, Alexis Diamond and Jens Hainmueller. 2010. "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program." *Journal of the American Statistical Association* 105(490):493–505.

Abadie, Alberto and Javier Gardeazabal. 2003. "The Economic Costs of Conflict: A Case Study of the Basque Country." *American Economic Review* 93(1):113–132.

Acemoglu, Daron, Simon Johnson, Amir Kermani, James Kwak and Todd Mitton. 2016. "The Value of Connections in Turbulent Times: Evidence from the United States." *Journal of Financial Economics* 121(2):368–391.

Alesina, Alberto and Ekaterina Zhuravskaya. 2011. "Segregation and the Quality of Government in a Cross Section of Countries." *The American Economic Review* 101(5):1872–1911.

Alesina, Alberto, Stelios Michalopoulos and Elias Papaioannou. 2016. "Ethnic Inequality." *Journal of Political Economy* 124(2):428–488.

Atack, Jeremy, Fred Bateman, Michael Haines and Robert A. Margo. 2010. "Did Railroads Induce or Follow Economic Growth?: Urbanization and Population Growth in the American Midwest, 1850-1860." *Social Science History* 34(2):171–197.

Banerjee, Abhijit, Esther Duflo and Nancy Qian. 2012. "On the Road: Access to Transportation Infrastructure and Economic Growth in China." National Bureau of Economic Research Working Paper 17897.

Bazzi, Samuel and Christopher Blattman. 2014. "Economic Shocks and Conflict: Evidence from Commodity Prices." *American Economic Journal: Macroeconomics* 6(4):1–38.

Beck, Thorsten, George Clarke, Alberto Groff, Philip Keefer and Patrick Walsh. 2001. "New Tools in Comparative Political Economy: The Database of Political Institutions (DPI)." *The World Bank Economic Review* 15(1):165–176. Dataset (url).

Berman, Eli, Jacob N. Shapiro and Joseph H. Felter. 2011. "Can Hearts and Minds Be Bought? The Economics of Counterinsurgency in Iraq." *Journal of Political Economy* 119(4):766–819.

Berman, Nicolas, Mathieu Couttenier, Dominic Rohner and Mathias Thoenig. 2017. "This Mine Is Mine! How Minerals Fuel Conflicts in Africa." *American Economic Review* 107(6):1564–1610.

Besley, Timothy and Torsten Persson. 2011. "The Logic of Political Violence." *The Quarterly Journal of Economics* 126(3):1411–1445.

Besley, Timothy, Torsten Persson and Daniel M. Sturm. 2010. "Political Competition, Policy and Growth: Theory and Evidence from the US." *The Review of Economic Studies* 77(4):1329–1352.

Cavallo, Eduardo, Sebastian Galiani, Ilan Noy and Juan Pantano. 2013. "Catastrophic Natural Disasters and Economic Growth." *Review of Economics and Statistics* 95(5):1549–1561.

Chandra, Amitabh and Eric Thompson. 2000. "Does Public Infrastructure Affect Economic Activity?: Evidence from the Rural Interstate Highway System." *Regional Science and Urban Economics* 30(4):457–490.

Colella, Fabrizio, Rafael Lalive, Seyhun Orcan Sakalli and Mathias Thoenig. 2019. "Inference with Arbitrary Clustering." IZA Discussion Paper n. 12584.

Desmet, Klaus, Ignacio Ortuño-Ortín and Romain Wacziarg. 2012. "The Political Economy of Linguistic Cleavages." *Journal of Development Economics* 97(2):322–338.

Desmet, Klaus, Joseph Flavian Gomes and Ignacio Ortuño-Ortín. 2019. "The Geography of Linguistic Diversity and the Provision of Public Goods." *Forthcoming in the Journal of Development Economics* .

Dinkelman, Taryn. 2011. "The Effects of Rural Electrification on Employment: New Evidence from South Africa." *American Economic Review* 101(7):3078–3108.

Duflo, Esther and Rohini Pande. 2007. "Dams." *The Quarterly Journal of Economics* 122(2):601–646.

Eberle, Ulrich J., Dominic Rohner and Mathias Thoenig. 2019. "Clashes over Culture or Cattle? Mixed Settlement and Ethno-Economic Conflict in Africa.". Unpublished.

Eberle, Ulrich J., J. Vernon Henderson, Dominic Rohner and Kurt Schmidheiny. 2020. "Ethno-Linguistic Diversity and Urban Agglomeration." *Forthcoming in the Proceedings of the National Academy of Sciences* .

Esteban, Joan, Laura Mayoral and Debraj Ray. 2012. "Ethnicity and Conflict: An Empirical Study." *American Economic Review* 102(4):1310–42.

Exchange on Environment Conflict and Cooperation. 2019. "Conflict Factsheet.". Online Catalogue (url).

Fearon, James D. 2005. "Primary Commodity Exports and Civil War." *Journal of Conflict Resolution* 49(4):483–507.

Feenstra, Robert C., Robert Inklaar and Marcel P. Timmer. 2015. "The Next Generation of the Penn World Table." *American Economic Review* 105(10):3150–82. Dataset (url).

Financial Times. 2020. "The Mekong Delta: An Unsettling Portrait of Coastal Collapse.". Retrieved January 5, 2019, from (url).

Food and Agriculture Organization, United Nations. 2013. "FAOSTAT Database.". Dataset (url).

Grill, Günther, Bernhard Lehner, Alexander E. Lumsdon, Graham K. MacDonald, Christiane Zarfl and Catherine Reidy Liermann. 2015. "An Index-Based Framework for Assessing Patterns and Trends in River Fragmentation and Flow Regulation by Global Dams at Multiple Scales." *Environmental Research Letters* 10(1):015001.

Grossman, Guy, Jan H Pierskalla and Emma Boswell Dean. 2017. "Government Fragmentation and Public Goods Provision." *The Journal of Politics* 79(3):823–840.

Harari, Mariaflavia and Eliana La Ferrara. 2018. "Conflict, Climate, and Cells: A Disaggregated Analysis." *Review of Economics and Statistics* 100(4):594–608.

Henderson, J. Vernon, Adam Storeygard and David N. Weil. 2012. "Measuring Economic Growth from Outer Space." *The American Economic Review* 102(2):994–1028.

Henderson, J. Vernon, Tim Squires, Adam Storeygard and David N. Weil. 2017. "The Global Distribution of Economic Activity: Nature, History, and the Role of Trade." *The Quarterly Journal of Economics* 133(1):357–406.

Henisz, Witold J. 2002. "The Political Constraint Index (POLCON) Dataset.". Dataset url).

Hodler, Roland and Paul A. Raschky. 2014. "Regional Favoritism." *The Quarterly Journal of Economics* 129(2):995–1033.

ICTA-UAB. 2019. "Environmental Justice Atlas.". Online Catalogue (url).

Imbens, Guido and Joshua Angrist. 1994. "Identification and Estimation of Local Average Treatment Effects." *Econometrica* 62(2):467–75.

König, Michael D., Dominic Rohner, Mathias Thoenig and Fabrizio Zilibotti. 2017. "Networks in conflict: Theory and evidence from the great war of africa." *Econometrica* 85(4):1093–1132.

Latham, John, Renato Cumani, Ilaria Rosati and Mario Bloise. 2014. "Global Land Cover Share (GLC-SHARE) Database Beta-Release Version 1.0-2014.". Dataset url).

Lawson, Robert A., Ryan H. Murphy and Claudia R. Williamson. 2016. "The Relationship between Income, Economic freedom, and BMI." *Public Health* 134:18–25. Dataset (url).

Lehner, Bernhard, Catherine Reidy Liermann, Carmen Revenga, Charles Vörösmarty, Balazs Fekete, Philippe Crouzet, Petra Döll, Marcel Endejan, Karen Frenken, Jun Magome, Christer Nilsson, James C. Robertson, Raimund Rödel, Nikolai Sindorf and Dominik

Wisser. 2011. "High-Resolution Mapping of the World's Reservoirs and Dams for Sustainable River-Flow Management (GRanD v1.3)." *Frontiers in Ecology and the Environment* 9(9):494–502. Dataset (url).

Lehner, Bernhard and Günther Grill. 2013. "Global River Hydrography and Network Routing: Baseline Data and New Approaches to Study the World's Large River Systems." *Hydrological Processes* 27(15):2171–2186. Dataset (url).

Lewis, M Paul, Gary F Simons and Charles D Fennig. 2018. "Ethnologue: Languages of the World, nineteenth ed. SIL International, Dallas.".

Lipscomb, Molly and Ahmed Mushfiq Mobarak. 2011. "Decentralization and the Political Economy of Water Pollution: Evidence from the Re-Drawing of County Borders in Brazil." *Forthcoming in the Review of Economics Studies* .

Lipscomb, Molly, Ahmed Mushfiq Mobarak and Tania Barham. 2013. "Development Effects of Electrification: Evidence from the Topographic Placement of Hydropower Plants in Brazil." *American Economic Journal: Applied Economics* 5(2):200–231.

Marshall, Monty G., Ted R. Gurr and Keith Jaggers. 2012. "Political Regime Characteristics and Transitions, 1800-2008, Polity IV Project.". Dataset (url).

Michaels, Guy. 2008. "The Effect of Trade on the Demand for Skill: Evidence from the Interstate Highway System." *The Review of Economics and Statistics* 90(4):683–701.

Michalopoulos, Stelios and Elias Papaioannou. 2013. "Pre-Colonial Ethnic Institutions and Contemporary African Development." *Econometrica* 81(1):113–152.

Michalopoulos, Stelios and Elias Papaioannou. 2014. "National Institutions and Subnational Development in Africa." *The Quarterly Journal of Economics* 129(1):151–213.

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

Montalvo, Jose G. and Marta Reynal-Querol. 2005. “Ethnic Diversity and Economic Development.” *Journal of Development Economics* 76(2):293–323.

National Geographic. 2019. “Two-Thirds of the Longest Rivers No Longer Flow Freely–And It’s Harming Us.”. Retrieved July 14, 2017, from (url).

National Oceanic and Atmospheric Administration. 2013. “Version 4 DMSP-OLS Nighttime Lights Time Series.”. Dataset (url).

Nickell, Stephen J. 1981. “Biases in Dynamic Models with Fixed Effects.” *Econometrica* 49(6):1417–1426.

Nordhaus, William D. 2006. “Geography and Macroeconomics: New Data and New Findings.” *Proceedings of the National Academy of Sciences* 103(10):3510–3517. Dataset (url).

Nunn, Nathan and Nancy Qian. 2014. “US Food Aid and Civil Conflict.” *American Economic Review* 104(6):1630–1666.

Pesaresi, Martino, Aneta Florczyk, Marcello Schiavina, Michele Melchiorri and Luca Maffenini. 2019. “GHS Settlement Grid, Updated and Refined REGIO Model 2014 in Application to GHS-BUILT R2018A and GHS-POP R2019A, Multitemporal (1975-1990-2000-2015), R2019A.” European Commission, Joint Research Centre (JRC). Dataset (url).

Rohner, Dominic and Alessandro Saia. 2019. “Education and Conflict: Evidence from a Policy Experiment in Indonesia.” CEPR Discussion Paper DP13509.

Schiavina, Marcello, Sergio Freire and Kytt Mac Manus. 2019. “GHS population grid multitemporal (1975, 1990, 2000, 2015) R2019A.” European Commission, Joint Research Centre (JRC). Dataset (url).

Storeygard, Adam. 2016. “Farther on Down the Road: Transport Costs, Trade and Urban Growth in Sub-Saharan Africa.” *The Review of Economic Studies* 83(3):1263–1295.

Strobl, Eric and Rorbert O. Strobl. 2011. “The distributional Impact of large Dams: Evidence from Cropland Productivity in Africa.” *Journal of Development Economics* 96(2):432–450.

Sundberg, Ralph and Erik Melander. 2013. "Introducing the UCDP Georeferenced Event Dataset." *Journal of Peace Research* 50(4):523–532. Dataset (url).

Swee, Eik L. 2015. "Together or Separate? Post-Conflict Partition, Ethnic Homogenization, and the Provision of Public Schooling." *Journal of Public Economics* 128:1–15.

The Economist. 2016. "The Mekong - Requiem for a River.". Retrieved July 14, 2017 (url).

The Economist. 2019. "A Dam Threatens One of the World's Oldest Settlements.". Retrieved October 26, 2019 (url).

Vermote, E., C. Justice, I. Csiszar, J. Eidenshink, R. Myneni, F. Baret, E. Masuoka, R. Wolfe and M. Claverie. 2014. "NOAA Climate Data Record (CDR) of Normalized Difference Vegetation Index (NDVI), Version 4.". Dataset (url).

Wolf, Aaron T. 1999. "The Transboundary Freshwater Dispute Database Project." *Water International* 24(2):160–163.

World Bank. 2019. "World Development Indicators.". Dataset (url).

World Commission on Dams. 2000. *Dams and Development: A New Framework for Decision-making*. London: EarthScan Publications.

Zarfl, Christiane, Alexander E. Lumsdon, Jürgen Berlekamp, Laura Tydecks and Klement Tockner. 2015. "A Global Boom in Hydropower Dam Construction." *Aquatic Sciences* 77(1):161–170.

CENTRE FOR ECONOMIC PERFORMANCE
Recent Discussion Papers

1693	Abel Brodeur Andrew E. Clark Sarah Flèche Nattavudh Powdthavee	COVID-19, Lockdowns and Well-Being: Evidence from Google Trends
1692	Fabrice Defever José-Daniel Reyes Alejandro Riaño Gonzalo Varela	All These Worlds are Yours, Except India: The Effectiveness of Cash Subsidies to Export in Nepal
1691	Adam Altmejd Andrés Barrios-Fernández Marin Drlje Joshua Goodman Michael Hurwitz Dejan Kovac Christine Mulhern Christopher Neilson Jonathan Smith	O Brother, Where Start Thou? Sibling Spillovers on College and Major Choice in Four Countries
1690	Michael Amior Alan Manning	Monopsony and the Wage Effects of Migration
1689	Frank Pisch	Managing Global Production: Theory and Evidence from Just-in-Time Supply Chains
1688	Barbara Petrongolo Maddalena Ronchi	A Survey of Gender Gaps through the Lens of the Industry Structure and Local Labor Markets
1687	Nick Jacob Giordano Mion	On the Productivity Advantage of Cities
1686	Andrew E. Clark Anthony Lepinteur	A Natural Experiment on Job Insecurity and Fertility in France

1685	Richard Disney John Gathergood Stephen Machin Matteo Sandi	Does Homeownership Reduce Crime? A Radical Housing Reform in Britain
1684	Philippe Aghion Roland Bénabou Ralf Martin Alexandra Roulet	Environmental Preferences and Technological Choices: Is Market Competition Clean or Dirty?
1683	Georg Graetz	Labor Demand in the Past, Present and Future
1682	Rita Cappariello Sebastian Franco-Bedoya Vanessa Gunnella Gianmarco Ottaviano	Rising Protectionism and Global Value Chains: Quantifying the General Equilibrium Effects
1681	Felipe Carozzi Christian Hilber Xiaolun Yu	On the Economic Impacts of Mortgage Credit Expansion Policies: Evidence from Help to Buy
1680	Paul Frijters Christian Krekel Aydogan Ulker	Machiavelli Versus Concave Utility Functions: Should Bads Be Spread Out Or Concentrated?
1679	Antoine Dechezleprêtre David Hémous Morten Olsen Carlo Zanella	Automating Labor: Evidence from Firm-Level Patent Data
1678	Michael Amiior	The Contribution of Immigration to Local Labor Market Adjustment
1677	Swati Dhingra Silvana Tenreyro	The Rise of Agribusiness and the Distributional Consequences of Policies on Intermediated Trade

The Centre for Economic Performance Publications Unit
Tel: +44 (0)20 7955 7673 Email info@cep.lse.ac.uk
Website: http://cep.lse.ac.uk Twitter: @CEP_LSE