Chapter 1:
The Impact of Broadband Internet on Immunizations

Carrie Huffaker
Università Commmerciale Luigi Bocconi

Adviser: Dr. Diego Ubfal

October 6, 2017

1 Introduction

Vaccination rates are rising in the United States- but not everywhere. In many local communities, coverage rates for different vaccine-preventable diseases have fallen before the herd immunity threshold, leading to new outbreaks. Measles ceased to be *endemic* (with a transmission rate 12 months or longer) in 2001, yet outbreaks not only continue to occur but are worsening over time. Seventy-seven percent of outbreaks with 20 or more cases have occurred since 2010 (Clemmons et al., 2017).

Several explanations have arisen, including the fact that most parents lack first-hand experience of the diseases that vaccines protect against, and the increase in the required number of shots over time. An alternative explanation is internet use. Since internet access spread to the American public in the early- to mid-1990s, a profusion of vaccine-critical websites have arisen and acquired hold of a vast audience. These websites challenge medical authority, and encourage parents to decide for themselves whether or not to vaccinate their child. There is substantial evidence that parents are turning to the internet for health information and news. Is it possible that there is a relationship between falling local vaccination rates and exposure to the internet? Or has the internet made parents more informed and encouraged vaccine uptake?

In this paper, I explore the relationship between broadband internet access and vaccination rates in U.S. counties from 1999 to 2008. Drawing upon internet access data from the Federal Communications Commission and immunization data from the Centers for Disease Control, I show that internet access has not led to a decrease in vaccination rates. Indeed, those counties with a larger number of broadband dividers have seen an increase in the 4:3:1 series (4 doses of DTaP/DTP vaccine, 3 doses of polio vaccine and 1 dose of MMR vaccine). This expansion in coverage is largely due to the rising uptake of the polio vaccine that has resulted from increased internet exposure.

My paper relates to three strands of literature. The first is literature on impact of new media upon women's health, fertility and other social outcomes. Television has been shown to influence attitudes towards gender relations and childbearing across the developing world, by presenting alternative and idealized lifestyles for women that strongly contrast with those permissible under traditional social norms (La Ferrara et al., 2008; Jensen and Oster, 2009; La Ferrara, 2016). In the United States, the MTV program *16 and Pregnant* has reduced pregnancy among teenage viewers (Kearney and Levine, 2014).

The second strand of literature addresses how the diffusion of new technologies influences society, and political outcomes in particular. Stromberg (2004) found that during the Great Depression, localities with a higher proportion of radio listeners were allocated more relief funds by politicians. With the introduction of television, the opposite has occurred: Viewers have become less informed, and vote less (Gentzkow, 2006). When news reports provide voters with less information about local politics, they elect local politics elect politicians who are less responsive to constituent needs (Snyder and Stromberg, 2010). Finally, it has been shown that the introduction of Fox News shifted large numbers of voters to the Republican party (DellaVigna and Kaplan, 2007). In each of these cases, the key mechanism is *information*: The substitution effects generated by technological diffusion cause voters to become more or less informed, and so changes their political behavior.

The final strand addresses the impact of the diffusion of internet, in particular. Lelkes et al. (2015) have instead demonstrated that the internet has increased partisan hostilities. These papers also tend to focus upon political outcomes. Work on Germany and the U.K. has found a negative impact of internet access upon voter turnout (Falck et al., 2014; Gavazza et al., 2016), while Larcinese and Miner (2017) have found the opposite to hold true in the U.S.

The Larcinese and Miner paper is probably the closest to mine, as they study the same period of American history and use the same data. They adopt a different empirical strategy, however, and focus on voting and campaign contributions rather than on vaccine uptake. My contribution to the literature is to study the impact of technological change, and specifically of the internet, on a public health outcome like immunizations. While other papers have focused upon political repercussions of the expansion of internet access, many other aspects of American life have been transformed as a result. The information mechanism may act upon attitudes and preferences towards vaccination, as well. I show that the internet, like other forms of media, can impact choices about personal and family health.

The paper is structured in the following way. In the next three sections I provide background information on the development and growth of internet services in the American market. I discuss immunization requirements, and address how the internet has become a learning resource for health information and the impact it has had upon public attitudes towards vaccination. Section 5 presents the datasets and empirical strategy employed in the analysis. Estimation results and robustness checks are presented in section 5. The final section concludes with a discussion. All figures and tables are found at the end of the paper, together with a brief appendix.

2 The Internet

The development and diffusion of the internet in the U.S. occurred in three distinct phases. The *Advanced Research Projects Agency Network (ARPANET)* emerged in the late 1960s as an initiative of the Department of Defense to create a robust and redundant system of data transmission that could survive nuclear attack. Packet-switching provided the technology for transferring data between remote systems. The military origins of ARPANET had several implications. First, access to ARPANET was limited to defense agencies and to the researchers at university computer science departments that won contracts to create the network. Second, all projects seeking funding and approval had to be framed in terms of Department of Defense objectives. Third, the DOD managed and controlled both the network and its community of users.

The network's civilian designers envisioned ARPANET as a tool for resource sharing: using remote hardware and software to run programs and perform research. However, as the network grew novel applications of its capabilities arose. File sharing and e-mail both proved become wildly popular, with e-mail in particular generating the greatest volume of traffic on the network. This marked a fundamental shift in focus for the network, from providing access to remote systems and processing capacities to providing access to people and their ideas (Abbate, 1999). At the same time, personal computers emerged as a new market- the first PC, the Altair 8800, launched in 1975- which developed in symbiosis with ARPANET.

In 1983, military users were split off from ARPANET to form MILNET. ARPANET was redefined as a civilian research network in this second phase. Access was expanded beyond DOD contractors to encompass other types of researchers and networks. Local area networks (LANs) connecting universities (and businesses) became widespread throughout the 1980s. Research networks were able to connect to ARPANET as long as they adopted the TCP/IP suite of standard protocols. The number of networks constituting the internet soared from only 15 networks in 1982 to more than 400 by 1986 (Abbate).

In the early 1980s, the National Science Foundation began planning a nationwide network that would evolve out of the ARPANET backbone and eventually, take it over in 1986 (ARPANET itself was decommissioned in 1990). The *NSFNET* was a three-tiered system consisting of, at the highest level, the NSFNET backbone; in the middle, regional networks; and below, universities and research centers. The transfer of the internet out of military control greatly expanded its reach across the scientific and academic community, but not far beyond. Much like its predecessor, the NSFNET had an Acceptable Use Policy that severely restricted how the internet could be used. Non-profit research and education were permitted; any type of commercial activity was not.

In the final phase of its development, use of the internet spread beyond academia to the American public. For this commercialization to occur, management of the network had to pass out of government hands and to the private sector. By the late 1980s, thriving market demand for computer networking services justified such a move. In 1988, the National Science Foundation transferred responsibility for the NSFNET backbone to Merit (the Michigan regional network), IBM and MCI. Beginning with the New York State Education and Research Network (NYSERNET), many other mid-level regional networks spun off for-profit branches. These were the first commercial Internet Service Providers (ISPs). Many local telephone companies would also enter the ISP market.

In 1991 the NSF elected to allow ISPs to take over the provision of all internet services, dismantling the old NSFNET backbone. The privatization of the internet was complete (Greenstein, 2015), and the network was now open to any use by any user. Before market demand could grow, however, the internet had to become *accessible* to the average American. To all but expert users, much of the internet was impenetrable; it was difficult to understand *what* information was available on the network, let alone *how* to access it. A series of innovations on two sides- the commercial and the technical- enabled the transformation of the internet into a usable product.

Beginning in the 1980s, applications like bulletin boards and newsgroups become widely popular, creating new communities of internet users and teaching them how to communicate online. Although grassroots networks provided some of these services, commercial online systems like Compuserve, America Online and Prodigy attracted far more users. The chatrooms and e-mail services created by these providers, as well as bulletin boards and newsgroups, were the means by which many Americans were introduced to the internet for the first time. As the network evolved into the 1990s, online providers established platforms of curated, proprietary content in order to usher in new users to the WWW.

This internet differed drastically from what we are familiar with today. First, users lacked the ability to freely create and publish their own content. Second, there was no single "internet," but a "balkanized world of electronic commerce" (Greenstein). Users could only access content to which they had paid subscriptions: a carry-over from the old world of print media. Different services were somewhat incompatible with each other; it was sometimes impossible to transfer messages across different e-mail systems.

The WorldWideWeb application was developed by researchers at CERN and first distributed in 1991. It revolutionized the internet by allowing for the transmission of many different forms of media. While the old internet was text-only, the WWW permitted the sharing of images and eventually, audio and video. Two years later, the first Web browsers appeared (Mosaic, soon to be overtaken in the market by Netscape Navigator). Finally, the emergence of search engines enabled users to find and access information with ease. Each of these three innovations contributed to making the internet transparent, and allowed adoption to spread far beyond the original market of technical experts and enthusiasts (Greenstein). Moreover, ordinary users could move beyond the guided, proprietary platforms of online services- most of which transformed into simple ISPs- to explore the internet for themselves, and eventually make their own contributions to it. Although the network had existed for over two decades by this point, it is only in the early- to mid-1990s that the internet spread to the wider American public, and so began to influence preferences and choices.

3 Immunizations

At the beginning of the last century, the infant mortality rate in the U.S. was 20%. Of children who survived infancy, one in five would die before age five. Many of these deaths would not occur today; they were attributable to infectious diseases that the population is now protected against by immunizations (Meckel, 2004). Over the course of the 20th century, vaccines have been credited with eradicating smallpox, diphtheria and polio from the U.S. population. They have drastically reduced the incidence of diseases that once infected hundreds of thousands of children each year, like measles, pertussis and hepatitis A. One study estimates that vaccination of the 2009 birth cohort will prevent 42,000 lives and 20 million cases of disease, and save billions of dollars (Fangjun et al., 2014).

Until the passing of the Vaccination Assistance Act by Congress in 1962, many parents were unable to vaccinate their children. The private practices or local health departments responsible for administering the vaccines often lacked the necessary infrastructure to do so, and in some cases families were forced to pay for them out of pocket. Following the passage of the Act and of later legislation, a nationwide immunization program was created to fund state and local health departments and, under the guidance of Public Health Advisers from the Centers for Disease Control (CDC), work towards the goal of universal immunization (Hinman et al., 2011).

The CDC is also responsible for drawing up the recommended immunization schedule. Over the last generation, many new vaccines have been developed, tested and added to the schedule, including those protecting against hepatitis A and varicella (both licensed in 1995), pneumococcal conjugate (2000), influenza (2003), meningococcal conjugate (2005), rotavirus, and human papillomavirus (2006). The recommended regime now includes vaccines against 14 different diseases, or as many as 27 shots by 24 months of age (Hinman et al.).

Underlying mass vaccination campaigns is the concept of *herd immunity*, defined by the CDC as "a situation in which a sufficient proportion of a population is immune to an infectious disease (through vaccination and/or prior illness) to make its spread from person to person unlikely" (CDC, 2015). Herd immunity depends upon a number of factors, including the effectiveness of the vaccine against disease transmission; the *reproduction number* R_0 (number of secondary cases generated by a typical infectious individual); and the degree to which vaccinated individuals are distributed randomly throughout the population (Fine et al., 2011). Assuming a completely effective vaccine and a randomly mixing, homogeneous population, the critical vaccination level q_c necessary to achieve herd immunity can be calculated as:

$q_c = 1 - \frac{1}{R_0}$

The herd immunity thresholds for six different vaccine-preventable diseases are presented in Table 3. More infectious diseases, likes measles and pertussis, have much higher thresholds.

If a vaccine is shown not to be effective; if unvaccinated individuals are not randomly distributed, but concentrated in specific communities, then herd immunity will fail and populations will become susceptible to disease. While the first risk lies within the realm of medical knowledge and expertise, the second is far more problematic to identify and address.

When immunization rates drop low enough in a given locality, diseases easily circulate through the affected population; eradicated diseases may even return to being *endemic* in the population: With a transmission rate lasting 12 months or longer (Clemmons et al., 2017). In the 1970s, a vaccine scare caused the government to abandon the vaccine, and the pertussis coverage rate to fall from 90% to 12% in Sweden. A pertussis epidemic broke out soon after, causing 10,000 cases of pertussis in a single year and many deaths (ECDC, 2012). In 1998, *The Lancet* published a paper linking the MMR vaccine to rising rates of autism. Although the paper has since been retracted, and the author lost his medical license, the ideas it put forth have had lasting consequences for MMR coverage rates. Ireland experienced a measles outbreak, with more than 300 cases and three deaths (Poland and Jacobson, 2011).

In the U.S., national immunization rates have continued to increase over time. This achievement masks important disparities at the local level, however. Herd immunity fails for many of the counties in my dataset. Comparing the average vaccination rate over the 1999-2008 period to the upper and lower thresholds for herd immunity against the six diseases presented in Table 3, it is apparent that all counties are fully immunized against polio, mumps and rubella (Figure 1). However, a large proportion are vulnerable to diptheria, measles and pertussis. A third of counties fall below the lower boundary for diptheria, while an additional 29% fall below the upper boundary. Only 40% of counties are protected. Forty-six percent of counties fall below the lower boundary for measles; 46% more, below the upper boundary, leaving only 7% of counties that are protected. Pertussis represents the gravest cause for concern. Only 4% of counties are adequately immunized. Fully 96% fall below the lower boundary for herd immunity.

This disparity in vaccination levels across U.S. counties reflects two important sets of factors. First, vaccination requirements are not enforced at the national level. The CDC recommended schedule is a guideline for parents and health care professionals, not a legal requirement. The responsibility for establishing the number of vaccines required of children in order to be admitted to childcare, kindergarten and 7th grade lies with state legislatures. Local school districts may choose to enforce these laws by turning away unvaccinated children, or look the other way; the degree of discretion enjoyed by individual institutions is somewhat limited, however, by the requirement that each school submit immunization statistics for the student body to the state health department at the beginning of the school year.

Vaccination requirements vary widely across states. Moreover, states may permit parents to opt of requirements by allowing for a variety of exemptions. Medical exemptions are legal in all 50 states; survivors of childhood cancer and other individuals with compromised immune systems cannot be safely vaccinated, for example. Several religious groups, including Christian Scientists and the Amish, forbid vaccination. In 2016, every state save West Virginia and California allowed for this type of objection to be protected by religious exemption. Personal belief, or philosophical, exemptions were allowed in 20 states (NCSL, 2016).[1] Different rules apply during outbreaks of VPDs; students enjoying such exemptions are usually required to remain at home for the duration of the outbreak, though this restriction is frequently challenged in court.

The second set of factors shaping idiosyncratic rates of vaccine uptake across localities are county-specific resources, preferences, and the information set available to individuals in the community in making the decision whether to vaccinate their children. As we will see in the next section, the advent of the internet revolutionized how people made decisions about their health. The quantity and quality of online content differs dramatically from the books, news reports and conversations with healthcare professionals that individuals may have relied upon in the past in gathering information.

4 The Internet and Immunizations

By the mid-1990s, the advent of the WorldWideWeb, Web browsers and efficient search engines had enabled the American public to access the internet at far higher rates than every before. Privatization of the internet backbone led to the growth of many new ISPs to satisfy the soaring demand. This expansion in usership was eventually followed by an expansion in who was able to create and publish new content on the internet. Blog-publishing tools such as Blogger had appeared by the late 1990s, enabling the non-technical public to build their own websites for the first time. New forms of social media followed, including Facebook (2004), Youtube (2005) and Twitter (2006). The explosion of user-generated content, together with the near-universal online presence of "real-world" organizations and institutions, has produced a democratization of information without precedent.

4.1 The New Information Economy

Studies carried out by the Pew Research Center reveal the increasing dependence of Americans upon the internet as a source of information. By 2000, 55% of Americans with internet access had looked up health information online; within two years, this figure had risen to 80% (Fox and Rainie, 2000;

[1] California allowed all three types of exemptions prior to the 2016-2017 school year. The December 2014 measles outbreak- which originated at the Disneyland Resort in Anaheim, CA- prompted the state senate to outlaw both religious and personal belief exemptions (SB277). Medical exemptions have risen since the passage of the law. Interestingly, the largest increase has occurred in those counties with a history of high PBE use (Delamater et al., 2017).

Fox, 2005). By 2006, a large proportion of the population had also come to rely upon the internet as their primary source of science news and information (Horrigan, 2006). The science study is interesting for two reasons. First, it reveals that individuals with home broadband connections use the internet differently. They are far more likely than dial-up users to obtain most of their science information from the internet (34% vs. 22%). This result also holds in the context of health: 88% of broadband users researched health information online in 2009, versus 72% of dial-up users (Fox and Jones, 2009). Broadband access is associated with a substitution away from other sources of information like television, magazines and newspapers.

Second, young adults age 18-29 (followed closely by adults age 30-49), are more likely than any other demographic to look up health and science topics on the internet, and to use it as their primary source of information. This reliance has increased over time. The same result holds true for parents of children under age 18. The implications are clear: many young parents learn about health and science online. In many cases- about four in ten- web resources are their principal learning reference. It is possible that these habits may extend to exploiting online resources when deciding whether or not to vaccinate children.

The main explanation for this reliance on the internet, given by 7 of 10 online seekers, is convenience (Horrigan). The internet has drastically lowered search costs, creating a world in which information is cheap and plentiful. Yet online research is characterized by one major drawback: a wealth of unsubstantiated and false "facts" compete with good sources in the online marketplace. Three-quarters of online seekers do not consistently check the date and source of online health information (Fox and Jones). One in four feel overwhelmed by the sheer amount of available health information, while one in five feel confused by what they read online. Similarly, a Pew report on internet use during the 2008 electoral campaign found that it is "usually difficult" for 56% of online adults to distinguish between what is true and what is false on the internet (Smith, 2011).

The online "antivax" movement has exploited this ambiguity between true and false information in creating an audience. Content analyses of prominent anti-vaccine websites- most of which present themselves as objective explorations of truth- reveal a common narrative placing doctors and other authorities in opposition to concerned parents. Equating the decision to vaccinate with passive, ignorant parenting, and a vaccine-skeptical stance as personal autonomy and empowerment for parents, rejection of the scientific evidence supporting vaccines is seen as "informed choice." The movement encourages parents to assume the role of self-taught experts on vaccinations and their child's health (Kata, 2010). It also accuses the medical establishment of hiding information about the negative consequences of vaccines in order to maximize profits for themselves and for pharmaceutical companies (ECDC, 2012; Betsch et al., 2012). Anti-vaccine websites frequently reject statistics and scientific evidence- indeed, the scientific process as a whole- on the grounds that the true risks of vaccination are subjective, unknown, and ultimately *unknowable* (Hobson-West, 2007; Brownlie and Howson, 2005).

Experimental evidence suggests that anti-vaccine websites can influence real-world vaccine decisions through changes in the preference for vaccination. Viewing vaccine-critical content increases the perceived risk of vaccines, while decreasing the perceived risk of not vaccinating. The perceived benefit of vaccines also declines. The impact of the personal narrative evidence frequently characterizing such websites is especially strong. Overall, there is a significant connection between exposure to anti-vaccine websites and reduced vaccine intentions (Betsch et al., 2010; Betsch et al., 2011; Kata).

Indeed, beliefs are fundamental to vaccination behavior. A significant share of parents who choose to delay or decline one or more required vaccines believe that vaccines cause both short-term side effects and potentially longer-term, more severe side effects. For these parents, following the CDC-recommended schedule is *more dangerous* than the alternative of no immunizations at all, or delayed ones. Such beliefs are widespread even among parents who adhere to the schedule (Dempsey et al., 2015). Vaccines are perceived, in many cases, as riskier than the diseases they prevent (Saada et al., 2015).

5 Data and Empirical Strategy

My analysis relies on a variety of different sources, including internet data from the Federal Communications Commission (FCC), immunization data from the CDC, and lightning density data from the National Centers for Environmental Information (NCEI). In this section I introduce the characteristics of each dataset, how they are collected, and the challenges created by data limitations. Table 1 presents descriptive statistics for all of the variables described. At the end of the section, I address the issue of identification and present the empirical strategy followed in section 6.

5.1 The Datasets

Internet Service Provision

In 1996, Congress enacted the first major change to telecommunications law since 1934. The aim of the Telecommunications Act was to remove regulatory barriers to entry and promote local competitiveness of the growing internet services industry. To this end, in March 2000 the FCC began to collect data from providers of *high-speed internet* (also termed "broadband"), defined as over 200 Kpbs in at least one direction. Until 2004, any ISP providing more than 250 high-speed connections in a state was required to report to the FCC; afterwards this provision was extended to all ISPs. The first four years of data thus underestimate the true number of high-speed service providers, particularly in rural areas.

The semi-annual FCC Broadband Progress Reports detailing the U.S. market for internet services between December 1999 and June 2008 provide the basis for my data analysis. In December 1999 ISPs supplied 2.8 million high-speed lines in the United States, including 1.8 million lines to residential and small business subscribers. By June 2000, this number had increased to 4.3 million HS lines, including 3.1 million to residential and small business subscriptions. One year later, this had grown to 9.6 million HS lines and 7.8 million residential and small business subscriptions (FCC, 2000a, 200b, 2002). By June 2008 there were 132.8 million lines in operation to homes and businesses (FCC, 2009).

Initially, however, the majority of internet subscriptions were not high-speed at all, but dial-up services that utilized the existing infrastructure of telephone lines. Interexchange Carriers, also known as long-distance phone companies, controlled the basic transmission mechanism for internet backbone traffic. Even more important, however, were the Local Exchange Carriers (LECs)- local telephone companies- which possessed the *last mile* of copper wire supplying telephony and internet services to nearly every home and business in the U.S. This monopoly granted the incumbent LECs a considerable amount of market power in the burgeoning market.

Newer technologies enabled data transmission at much higher bandwidths. Digital subscriber lines (DSL) provided increased bandwidth from the existing copper loops of local telephone companies, and unlike dial-up service did not interfere with the carriage of voice service. By 1998, many incumbent LECs had begun to offer DSL service. At the same time, the innovation of broadband access via coaxial cable enabled a new participant- local cable companies- to enter the market. This entrance was eased by the fact that 2/3 of American households already subscribed to cable television. As with LECs, local cable companies enjoyed market power through the ownership of the last mile of coaxial cable. This presented a challenge to outside ISPs. While many LECs and cable companies themselves entered the market as internet providers, outside ISPs depended on the last mile facilities of others for access to retail customers (FCC, 1998). Nonetheless the number of ISPs operating in the U.S. grew quickly over time.

Other sources of broadband service included satellite, wireline, and fixed wireless technologies (FCC, 2000a), although these represent a minority of total supply from 1999 to 2008. Most households accessed the internet through coaxial cable (93% of high-speed lines in 2000; 34.1% in 2008), even as DSL service grew much faster, from only 7% of high-speed lines in 2000 to 27.3% in 2008.

Limitations of the Internet Data

The FCC only collects data from high-speed internet providers, not the universe of providers as a whole. We thus lack information about the dial-up market. However, there is reason to focus on broadband rather than internet access in general. As discussed above in the context of health and science news, broadband access is associated with both a substitution away from other forms of media, and a change in how individuals use the internet. Not only do they spend more time online, but there is extensive evidence of behavioral differences in home broadband users. In addition to looking up information, they are more likely to share files and download media of all types (Horrigan and Rainie, 2002). Even more importantly, broadband users are an important source of user-generated content. To a far greater extent than dial-up users, they post personal content like artwork, photos and stories on the internet; they are also more prone to create their own websites (Horrigan, 2007). It is reasonable to predict that high-speed access will also have a differential impact upon off-line behavior like vaccinations.

A second limitation is the level at which the FCC collects and reports data. Up to June of 2008, the FCC only provided data at the state and zip-code level. Because the CDC immunization data is reported by U.S. county, it is necessary to aggregate the zip-code data up to the county level. I performed this operation using the Missouri Census Data Center's MABLE/Geocorr web application. Each zip-code observation was matched to one or more counties and then weighted by the proportion of county land area that it contributes. County-level figures for number of high-speed service providers therefore represent an extrapolation from the relative data of all constituent zip codes.

A third limitation is that the FCC only reports number of high-speed service providers at the zip-code level, ignoring the demand side of the market. It is common in the literature to employ number of providers as a proxy for the degree of internet penetration (see Larcinese and Miner, 2017, for one example). By this metric, every county in my dataset had internet access by 1999; yet the degree of penetration varies widely, from a minimum of 0.01 to a maximum of 20.38 providers (Table 1).

Finally, zip codes with one, two or three providers are grouped together into a single class. A Freedom of Information Act request for the actual number of providers in these zip codes was refused on the grounds that the information represented a trade secret, and release of it could cause substantial competitive harm (Exemption 4 to the Act). I follow the FCC in assigning a value of "1" to these zip codes.

Despite these limitations, there is a strong correspondence between the constructed county measures for the December 1999-June 2008 period and the county-level data reported by the FCC for subsequent periods. In December 2008, the FCC began to record data at the county level for five variables. Two of these address demand for high-speed services: Number of residential connections over 200 kbps in at least one direction per 1000 households, and number of residential connections at least 3 Mbps downstream and at least 768 kbps upstream per 1000 households.[2] The last three variables address supply. The total number of providers of connections over 2000 kpbs in at least one direction corresponds exactly to my provider measure. Additionally, the FCC reports the number of providers of residential high-speed connections and of mobile high-speed connections.

Tables 4 and 5 demonstrate the interrelationship between all of these different variables. Reports in Table 4 are the correlations between my constructed measure for June 2008 and the five FCC variables in December 2008, from 2008-2009 (an average of three observations), and from 2008-2011 (an average of six observations). The correlation between my calculations for June 2008 and the December 2008 data is quite high, equal to 0.80 (column (1)). As columns (2) and (3) show, this strong correlation persists across longer time periods. My measure is also correlated with number of residential providers (ρ= 0.62) and mobile providers (ρ= 0.50). The two lagged demand variables, as well, exhibit a positive relationship with number of providers. A larger market supply, associated perhaps with a more competitive local market for internet services and lower prices, is associated with higher demand for internet connections. The correlation between providers and connections grows even larger when "high-speed" is more strictly defined (ρ= 0.522).

Table 5 reports the same set of correlations, substituting the December 2008 FCC data for my constructed June measure. The relationships are stronger, yet of roughly the same magnitude and display a similar degree of persistence over time. The tables present compelling evidence of two facts. My calculations represent a good match to the true number of providers per county for the 1999-2008 period; and in all periods there is a positive relationship between the supply and demand sides of the market, which grows more significant when attention is restricted to connections of larger bandwidths and speeds.

Immunizations

[2]One kilobyte represents 1,024 bytes; one megabyte is 1,024 kilobytes. The second definition of "high-speed" is therefore much stricter.

My principal source of immunization data is a Surveillance Summary published by the CDC in 2011. The Summary is drawn from National Immunization Survey (NIS) records for the 1995-2008 period (in order to match my internet data, I focus only on 1999-2008). The NIS is a random-digit-dial telephone survey conducted by the CDC's National Center for Immunization and Respiratory Diseases. The survey originated in 1994 in response to a series of measles outbreaks earlier in the decade; its purpose is to monitor nationwide vaccination coverage among children and teens for a variety of infectious diseases. The Surveillance Summary focuses on the population of children aged 19-35 months. The sample is composed of 257 counties for which the NIS sample size exceeded 35 observations at least once in seven biennial periods. Given that the entire U.S.- including outlying territories- numbers 3,143 counties, it is clear that a relatively limited share of the most populous counties are included in the study. Table 2 presents descriptive statistics relative to the immunization data for this sample.

My main outcome of interest is the coverage rate for the 4:3:1 series, defined as 4 doses of DTaP/DTP vaccine, 3 doses of polio vaccine, 1 dose of MMR vaccine. The DTap/DTP vaccine protects against diptheria, tetanus and pertussis. The polio vaccine is administered alone. The MMR vaccine protects against measles, mumps and rubella. In addition to the 4:3:1 vaccination rate, I analyze separately the uptake of each of the three vaccines. I also include a number of other series for which data is available for the entire 1999-2008 period: The vaccines for Haemophilus influenzae type B (Hib), Hepatitus B, and varicella. Summary statistics for the three vaccines are presented in Table 3. On average, immunization rates were lowest for varicella and highest for Hib, although Hib coverage declined over the entire period. The greatest growth in coverage occurred for varicella.

Descriptive evidence for the relationship between degree of broadband penetration and vaccine uptake is provided by Figures 2-8, which trace the evolution of immunization behavior by number of broadband providers. Counties are divided into two classes according to the initial number of providers in 1999. The two classes appear fairly similar at the beginning of the period. Provider-rich counties (more than three ISPs) begin with lower coverage rates for three vaccines (polio, Hib and hepatitis B); higher coverage rates for MMR and varicella; and about the same rate for DTap/DTP and the 4:3:1 series. Yet by the end of the period, provider-rich counties have consistently higher rates for all series except two: Polio, which is nearly the same for the two groups, and hepatitis B, for which provider-poor counties have a higher coverage rate. These figures provide early evidence that immunization rates rose more quickly in those counties with a higher exposure to high-speed internet.

County Control Variables

Many of the same factors have been shown to be associated with both the supply of internet and local-level immunization rates. Two of the most important determinants of internet supply are population density and median family income. The ethnic make-up of a county is also important (FCC, 2000). Attitudes towards vaccines and vaccination behavior are strongly correlated with characteristics like income, ethnicity, age and level of education (Funk et al., 2017). The foreign-born population also matters for vaccination rates. Immigrant populations tend to exhibit systematically different rates of vaccination from native Americans, and to vaccinate their children at different rates.[3] Furthermore, diseases that are no longer endemic in the United States- like measles- continue to be imported from endemic countries by international travelers (Clemmons et al., 2017). Such localized outbreaks have been shown to influence immunization rates in turn, generating a feedback effect (Oster, 2017).

In this paper, I include controls for economic status (median household income, the poverty rate of families with dependent children under age 17, and the employment rate), education (percent of population with a college degree or higher), demographic make-up (percent foreign-born, black and Hispanic); and finally, for composition of employment by industry. The four industries included as controls are information; finance, insurance and real estate; professional, scientific, and management, and administrative and waste management services; and educational services, health care and social assistance. All controls are taken from the 1990 and 2000 Censuses, and the American Community Surveys of 2005-2008.

5.2 Identification and the Lightning Instrument

The effect of access to broadband internet upon vaccine uptake cannot be identified with OLS because the pattern of internet provision that we observe is unlikely to be exogenous. Among the earliest actors to adopt high-speed internet were educated individuals with a high degree of technological knowledge, who tended to be clustered in those parts of the country connected to the NSFNET backbone. If they had a different understanding and opinion of the benefits and risk of immunization, this would generate a spurious correlation between internet supply and immunization rates. Similarly, reverse causality could arise if families subscribed to broadband internet in order to be able to perform online research about health and science. These families might choose to connect to the internet because of a particular interest in vaccination and other health topics, not vice-versa.

Identification requires us to isolate that part of variation in internet supply that is exogenously determined. Three sorts of instruments for internet supply are common to the literature. The first are Right-of-Way laws, or the regulations enacted by U.S. states that determine how difficult it is for new firms to enter the ISP market. Larcinese and Miner (2017) and Lelkes et al. (2015) employ this instrument. It is difficult to prove that variation in these state laws is truly exogenous. A larger barrier is data availability. Each paper relies upon the 2002 Technet State Broadband Index, which classifies and evaluates legislation passed after my period of study has begun.[4]

Several studies instrument supply by exploiting variations in the pre-existing telephony infrastructure (Campante et al., 2013; Falck et al., 2014). This is particularly suitable for countries like Germany and Italy, where national telecom companies initially exercised a monopoly over the provision of internet service. Because American consumers had a larger range of options to choose from in selecting service providers, the instrument is inapplicable to the U.S. context.

Natural phenomena such as rainfall and lightning density represent a last set of potential in-

[3] A Somali-American community in Minnesota had among the highest vaccine coverage rates in the state until 2008, when a growing number of autism diagnoses caused concern about the MMR vaccine. The community entered into communication with the anti-vaccine movement, and coverage rates plummeted. As a consequence, the community experienced a serious outbreak of measles in May of 2017 (Howard, 2017).

[4] Microsoft Technet is a web service for IT professionals. It has published two State Broadband Indices, in 2002 and 2012.

struments, employed by Gavazza et al. (2016) and Andersen et al. (2011). Because flooding and lightning strikes damage electrical equipment and infrastructure, it becomes relatively more costly to supply internet services in affected localities. As the *Wall Street Journal* observes, “in the information age, lightning matters in ways it didn’t before. A blip in the power supply that in the 1980s might have caused only a flickering of the lights can now fry circuits, shut down servers and deaden phone systems” (Lahart, 2009). *The Economist* (2001) notes that millions of dollars are lost each year due to “seemingly insignificant power faults that cause assembly lines to freeze, computers to crash and networks to collapse.”

I use lightning count data collected by the U.S. Air Force 14th Weather Squadron in collaboration with Vaisala’s U.S. National Lightning Detection Network (NLDN), and made available to the public by the NCEI. The NLDN monitors cloud-to-ground lightning activity using ground-based sensors. These sensors detect the electromagnetic signals generated by lightning strikes and submit the data to the NLDN Network Control Center in Tucson, Arizona (Vaisala, 2015).

Identification relies upon an instrument that is relevant, exogenous and satisfies the exclusion restriction. Atmospheric phenomena such as lightning are exogenous in the absence of measurement error, and relevance is demonstrated by a strong first stage. I address the exclusion restriction by checking for balance on covariates of the lightning instrument. First, I divide the sample in half according to whether observations fall above or below the median value for lightning strikes. Characteristics of the “low lightning” class (with a mean density of lightning strikes per year less than 4356) are shown in column (1) of Table 6; those of the “high” class (density greater than 4356 strikes per year), in column (2). Column (3) presents the mean difference and associated level of significance.

The instrument is balanced on population, employment, and three of the four of industry indicators (employment in education is instead significant at the 10% level). Balance fails, however, for the household economic and demographic variables. This is not surprising. There is great heterogeneity in the distribution of lightning strikes across U.S. territory, with the highest density of observations concentrated in the coastal southeast and a relative paucity of activity along the Pacific Coast. By controlling for state and county fixed effects, we can account for geographic differences in the composition of the American population. In addition, it is important to control for all the statistically significant variables in the regression analysis.

I intend to perform an additional test to demonstrate that the exclusion restriction holds for the lightning instrument. Lightning density may influence the provision of public health services in a county, which in turn drives local variation in immunization rates. To account for this possibility, I will introduce as controls the number of hospital beds, pediatricians and doctors as a whole per 1,000 inhabitants, and check for balance on these variables as well.

5.3 Empirical Strategy

I test the hypothesis that broadband access influences vaccination rates by running a fixed-effects regression of the form:

$$v_{ijt} = \alpha_i + \gamma_j + \delta_t + \psi prov_{it} + \beta X_{it} + \varepsilon_{it}$$

where counties are indexed by i, states by j, and years by t. The vaccination rate for county i in state j in year t is given by v_{ijt}. In the empirical analysis I model each of the seven vaccines presented in Table 2. The treatment variable, number of internet providers, is $prov_{it}$. The vector of time-varying observables is given by X_{it}. All models include county, state and year fixed effects. Errors are clustered at the county level.

The identifying assumption is that in the absence of exposure to internet, counties experiencing a more intense form of treatment (in the form of more providers) and counties experiencing a less intense form of treatment (fewer providers) would have changed the same way. The intensity of treatment is the key factor varying across counties, given that all counties had at least one internet provider at the beginning of the study period. In other words, considerations common to all counties- such as the dramatic increase in the number of required vaccinations over time, which has placed a heavier burden upon parents; the invention of new and more complex formulations, which immunize children against many diseases at the same time; and the loss of living historical memory of most of the diseases vaccinations protect against- given all of these common considerations, provider-rich and provider-poor counties would have experienced the same rise and fall in vaccination rates. Conditional on observables, the driving force behind differential trends is exogenous variation in the intensity of the internet treatment, as determined by the lightning instrument.

6 Does Internet Access Drive Vaccination Behavior?

4:3:1 Series Results

I begin by presenting estimation results for the 4:3:1 series, followed by those for the six individual vaccines. Column (1) of Table 7 provides the OLS coefficient estimate for the univariate regression of vaccination coverage on the provider variable, while county covariates are added in column (2). In each case the relationship is positive and significant, though the level of significance falls when observables are added. The column (2) coefficient suggests that each additional provider in a county is associated with an increase of 0.12 points in the 4:3:1 coverage rate.

The IV model is presented in columns (3)-(6). The 1st-stage coefficient in column (3), relative to the univariate model, is positive and extremely significant. A higher density of lightning strikes is associated with *more* internet providers, in marked contrast to the finding of Betsch that lightning activity slows the diffusion of internet. The F-statistic for the regression is quite large, equal to 26.15. The coefficient for the univariate second stage, in column (4), is about 4.5 times larger in magnitude than the estimate obtained using OLS, demonstrating a clear negative bias in the OLS results. A one standard deviation increase in the intensity of internet access (4.13 new providers) causes the vaccination rate to rise by 2.75 points. This is considerable relative to the mean coverage rate for the entire period, equal to 82%.

The point estimate increases to 0.756 when covariates are added. In this case, a one standard deviation in the number of providers causes the coverage rate to rise by 3.12 points. With and without covariates, the estimated effect is significant at the 5% level.

Individual Vaccine Series Results

The OLS estimate results in columns (1) and (2) of Table 8 suggest a positive and highly significant relationship between internet access and the polio, Hib and hepatitis B vaccines. Evidence for a positive correlation is not entirely consistent, however: A much larger and negative relationship is observed for the varicella series.

The broadband effect is not robust once we instrument for number of providers. As columns (3) and (4) show, the Hib, hepatitis B and varicella estimates lose significance. Yet the coefficient on polio increases, and retains its significance at the 1% level. Comparing the results of Tables (3) and (4), it appears that the positive relationship between internet access and the 4:3:1 series is almost entirely driven by increased uptake of the polio vaccine.

7 Robustness Checks

In this section, I will perform a number of different robustness checks. First, I will test whether variation in the lightning instrument is predictable from initial vaccination rates (in the year 1995). This is an another test of the exclusion restriction. It is important to verify whether there is some unobservable that is correlated with lightning and impacts the outcome of interest.

Second, I will interact initial vaccination rates with the year variable in order to create a time trend for each county. By controlling for initial conditions, I verify that the estimated treatment effect is not due to pre-existing differences in vaccination rates.

Third, I will introduce into the model historical data on outbreaks of vaccine-preventable diseases. One common explanation for changing local vaccination rates is that most parents lack first-hand experience with these diseases, causing them to systematically underestimate the risk of vaccine refusal. Counties that experienced more severe outbreaks in the past may be experiencing an increase in vaccine coverage, while those counties less exposed to disease may be undergoing a decline. By accounting for different patterns of disease in the past, I can discount the "historical memory" argument.

Fourth, I will introduce dummy variables for religious exemptions and personal belief exemptions. If allowable exemptions are changing over time across states, then state fixed effects will fail to capture their impact. The exemption dummies will ensure that changes in vaccine rates are not attributable to changes in state laws.

Finally, states such as California and Massachusetts have gathered county-level vaccination data dating back to the early 1990s: Before internet use had spread from the expert population to the American public. Using data from one or more of these smaller datasets, I can check for pre-existing trends in immunizations. I can exploit variation in internet access on the extensive rather than the intensive margin, given that many of the smaller counties will still lack access in 1999 (the first year for which I have local-level internet data). Furthermore, because such data is available up to the present, I can analyze vaccination trends over a much longer period of time. This analysis will enjoy much higher internal validity at the cost of some external validity (California and Massachusetts are both relatively wealthy, left-leaning, and hardly representative of the United States as a whole).

8 Discussion

In this paper, I find evidence that broadband access drives vaccination behavior. Greater exposure to the internet causes vaccination rates to rise. Counties with more broadband providers had higher 4:3:1 coverage rates, driven by increased uptake of the polio vaccine. The OLS model suggested a positive relationship between internet and the Hib and hepatitis B vaccines, and a negative one with varicella, but these results were not robust.

Mesch and Schwirian (2014) have linked perceptions about vaccines and vaccine intentions to confidence in the medical establishment and in government. Many participants in the antivaccine movement do not view these institutions as benign. Indeed, individuals who reject vaccines and the scientific worldview are more likely to hold conspiracy theory beliefs (Jolley and Douglas, 2014; Lewandowsky et al., 2013). Even among the public as a whole, however, confidence in the medical profession has declined over time (Schlesinger, 2002). Moreover, the American people hold starkly different beliefs from experts on a number of medical, scientific and technological issues- without even being aware of it (Funk, 2015).

The authority of scientists and doctors in society is in steady decline since peaking in the 1960s. Some experts argue that the decline has accelerated in recent decades, and while a number of factors are cited, the internet is always prominent among them (Nichols, 2017; Otto, 2016). As Nichols argues in the *End of Expertise*, "the Internet has accelerated the collapse of communication between experts and laypeople by offering an apparent shortcut to erudition. It allows people to mimic intellectual accomplishment by indulging in an illusion of expertise provided by a limitless supply of facts."

A positive relationship between broadband access and immunizations is thus both interesting and unexpected. It mirrors the finding of Larcinese et al.- after several other papers had uncovered a negative relationship with political variables- that the internet had increased voter turnout in the U.S. In all of these papers, the role of information is fundamental: Individuals are voting less, and politicians are becoming less accountable to their constituents, because substitution between different forms of media is rendering voters less informed. It is difficult to contextualize my findings in this literature- as far as I am aware, no other paper has studied the relationship between broadband and vaccination rates- but it is clear that in some contexts, the internet may have the converse effect of enabling the public to become more educated and informed. The extent to which public beliefs about science and the value of expert authority are also being transformed could be the subject of future research, together with the impact of internet upon other health outcomes.

References

[1] Introduction

[2] Abbate, J. 1999. *Inventing the Internet*. Cambridge: MIT Press.

[3] Andersen, T.B., J. Bentzen, C. Dalgaard and P. Selaya. 2011. "Does the Internet Reduce Corruption? Evidence From U.S. States and Across Countries." The World Bank Economic Review, 25(3): 387-417.

[4] Andersen, T.B., J. Bentzen, C. Dalgaard and P. Selaya. 2012. "Lightning, IT Diffusion, and Economic Growth Across U.S. States." *The Review of Economics and Statistics*, 94(4): 903-924.

[5] Betsch, C., N.T. Brewer, P. Brocard, P. Davies, W. Gaissmaier, N. Haase, J. Leask, F. Renkewitz, B. Renner, V.F. Reyna, C. Rossmann, K. Sachse, A. Schachinger, M. Siegrist and M. Stryk. 2012. "Opportunities and Challenges of Web 2.0 for Vaccination Decisions." *Vaccine*, 30(25): 3727-3733.

[6] Betsch, C., F. Renkewitz, T. Betsch and C. Ulshöfer. 2010. "The Influence of Vaccine-Critical Websites on Perceiving Vaccination Risks." *Journal of Health Psychology*, 15(3): 446-55.

[7] Betsch C., C. Ulshöfer, F. Renkewitz and T. Betsch. 2011. "The Influence of Narrative v. Statistical Information on Perceiving Vaccination Risks." *Medical Decision Making*, 31(5): 742-53

[8] Brownlie, J. and A. Howson. 2005. "Leaps of Faith and MMR: An Empirical Study." *Sociology*, 39(2): 221-239.

[9] Campante, F., R. Durante and F. Sobbrio. 2013. "Politics 2.0: The Multifaceted Effect of Broadband Internet and Political Participation." NBER Working Paper 19029.

[10] Centers for Disease Control and Prevention- National Center for Immunization and Respiratory Diseases. 2015. "Vaccines and Immunizations: Glossary." https://www.cdc.gov/vaccines/terms/glossary.html

[11] Clemmons, N.S., G.S. Wallace, M. Patel and P.A. Gastanaduy. 2017. "Incidence of Measles in the United States, 2001-2015." *Journal of the American Medical Association*, 318(3): 1279-1281.

[12] Delamater, P.D., T.F. Leslie and Y.T. Yang. 2017. "Change in Medical Exemptions for Immunization in California After Elimination of Personal Belief Exemptions." *Journal of the American Medical Association*, 318(9): 863-864.

[13] DellaVigna, S. and E. Kaplan. 2007. "The Fox News Effect: Media Bias and Voting." *The Quarterly Journal of Economics*, 122(3): 879–924.

[14] Dempsey, A.F., S. Schaffer, D. Singer, A. Butchart, M. Davis and G.L. Freed. 2011. "Alternative Vaccination Schedule Preferences Among Parents of Young Children." *Pediatrics*, 128(5): 848-856.

[15] European Center for Disease Prevention and Control. 2012. "Communication on Immunisation – Building Trust." Stockholm: ECDC.

[16] Fangjun, Z., A. Shefer, J. Wenger, M. Messonnier, L.Y. Wang, A. Lopez, M. Moore, T.V. Murphy, M. Cortese and L. Rodewald. 2014. "Economic Evaluation of the Routine Childhood Immunization Program in the United States, 2009." *Pediatrics*, 133(4): 577-585.

[17] Falck, O., R. Gold and S. Heblich. 2014. "E-lections: Voting Behavior and the Internet." *American Economic Review*, 104(7): 2238-2265.

[18] Federal Communications Commission. 1998. "First Broadband Progress Report."

[19] Federal Communications Commission. 2000a. "Federal Communications Commission Releases Data on High-Speed Services for Internet Access."

[20] Federal Communications Commission. 2000b. "Deployment of Advanced Telecommunications Capability: Second Report."

[21] Federal Communications Commission. 2002. "Third Broadband Progress Report."

[22] Federal Communications Commission. 2008. "Fifth Broadband Progress Report."

[23] Fine, P., K. Eames and D.L. Heymann. 2011. "Immunizations: A Rough Guide." *Clinical Infectious Diseases*, 52(7): 911-916.

[24] Fox, S. 2005. "Health Information Online." Pew Internet & American Life Project.

[25] Fox, S. 2006. "Online Health Search 2006." Pew Internet & American Life Project.

[26] Fox, S. and S. Jones. 2009. "The Social Life of Health Information." Pew Internet & American Life Project.

[27] Fox, S. and L. Rainie. 2000. "The Online Health Care Revolution." Pew Internet & American Life Project.

[28] Funk, C. 2015. "Public and Scientists' Views on Science and Society." Pew Research Center.

[29] Funk, C., B. Kennedy and M. Hefferon. 2017. "Vast Majority of Americans Say Benefits of Childhood Vaccines Outweigh Risks." Pew Research Center.

[30] Gavazza, A., M. Nardotto and T. Valletti. 2016. "Internet and Politics: Evidence from U.K. Local Elections and Local Government Policies."

[31] Gentzkow, M. 2006. "Television and Voter Turnout." *The Quarterly Journal of Economics*, 121(3): 931-972.

[32] Greenstein, S. 2015. *How the Internet Became Commercial*. Princeton: Princeton University Press.

[33] Hinman, A.R., W. A. Orenstein and A. Schuchat. 2011. "Vaccine-Preventable Diseases, Immunizations, and MMWR — 1961–2011." *Morbidity and Mortality Weekly Report Supplements*, 60(04): 49-57.

[34] Hobson-West, P. 2007. "'Trusting Blindly Can Be the Biggest Risk of all': Organised Resistance to Childhood Vaccination in the UK." *Sociology of Health and Illness*, 29(2):198-215.

[35] Horrigan, J.B. 2006. "Home Broadband Adoption 2006." Pew Internet & American Life Project.

[36] Horrigan, J.B. 2006. "The Internet as a Resource for News and Information about Science." Pew Internet & American Life Project.

[37] Horrigan, J.B. 2007. "Broadband: What's All the Fuss About?" Pew Internet & American Life Project.

[38] Horrigan, J.B. and L. Rainie. 2002. "The Broadband Difference: How Online Behavior Changes With High-speed Internet Connections." Pew Internet & American Life Project.

[39] Howard, J. 2017. "Anti-vaccine Groups Blamed in Minnesota Measles Outbreak." *CNN*, http://edition.cnn.com/2017/05/08/health/measles-minnesota-somali-anti-vaccine-bn/index.html.

[40] Jensen, R. and E. Oster. 2009. "The Power of TV: Cable Television and Women's Status in India." *The Quarterly Journal of Economics*, 124(3): 1057–1094.

[41] Jolley, D. and K.M. Douglas. 2014. "The Effects of Anti-Vaccine Conspiracy Theories on Vaccination Intentions." *Plos One*, 9(2).

[42] Kata, A. 2010. "A Postmodern Pandora's Box: Anti-Vaccination Misinformation on the Internet." *Vaccine*, 28: 1709-1716.

[43] Kearney, M.S. and P.B. Levine. 2014. "Media Influences on Social Outcomes: The Impact of MTV's 16 and Pregnant on Teen Childbearing." NBER Working Paper 19795.

[44] La Ferrara, E. 2016. "Mass Media and Social Change: Can We Use Television to Fight Poverty?" *Journal of the European Economic Association*, 14(4): 791-827.

[45] La Ferrara, E., A. Chong and S. Duryea. 2008. "Soap Operas and Fertility: Evidence from Brazil." Inter-American Development Bank Working Paper 633.

[46] Lahart, J. 2009. "There Go the Servers: Lightning's New Perils." *The Wall Street Journal*, March 22.

[47] Larcinese, V. and L. Miner. 2017. "The Political Impact of the Internet on U.S. Presidential Elections."

[48] Lelkes, Y., G. Sood and S. Iyenger. 2015. "The Hostile Audience: The Effect of Access to Broadband Internet on Partisan Effect." *American Journal of Political Science*, 61(1): 5-20.

[49] Lewandowsky, S., G.E. Gignac and K. Oberauer. 2013. "The Role of Conspiracist Ideation and Worldviews in Predicting Rejection of Science." *Plos One*, 8(10).

[50] Meckel, R.A. 2004. "Levels and Trends of Death and Disease in Childhood, 1620 to the Present." *Children and Youth in Sickness and Health: A Handbook and Guide*, ed. J. Golden, R.A. Meckel and H.M. Prescott. Westport, CT: Greenwood Press.

[51] Mesch, G.S. and K.P. Schwirian. 2014. "Confidence in Government and Vaccination Willingness in the U.S." *Health Promotion Internation*, 30(2): 213-221.

[52] National Conference of State Legislatures. 2016. "States with Religious and Philosophical Exemptions from State Immunization Requirements." http://www.ncsl.org/research/health/school-immunization-exemption-state-laws.aspx

[53] National Telecommunications and Information Administration. 1998. "Falling Through the Net II: New Data on the Digital Divide."

[54] Nichols, T. 2017. *The Death of Expertise*. Oxford: Oxford University Press.

[55] Oster, E. 2017. "Does Disease Cause Vaccination? Disease Outbreaks and Vaccination Response."

[56] Otto, S. 2016. *The War on Science.* Minneapolis: Milkweed Editions.

[57] Poland, G.A. and R.M. Jacobson. 2011. "The Age-Old Struggle Against the Antivaccinationists." *The New England Journal of Medicine*, 364(2): 97-99.

[58] Saada, A., T.A. Lieu, S.R. Morain, B.J. Zikmund-Fisher and E. Wittenberg. 2015. "Parents' Choices and Rationales for Alternative Vaccination Schedules: A Qualitative Study." *Clinical Pediatrics*, 54(3): 236-243.

[59] Smith, A. 2011. "The Internet and Campaign 2010." Pew Internet & American Life Project.

[60] Snyder, J.M. and D. Stromberg. 2010. "Press Coverage and Political Accountability." *Journal of Political Economy*, 118(2): 355-408.10/09/17

[61] Stromberg, D. "Radio's Impact on Public Spending." 2004. *The Quarterly Journal of Economics*, 119(1): 189–221.

[62] Vaisala. 2015. "Vaisala's NLDN® U.S. National Lightning Detection Network." https://www.vaisala.com/sites/default/files/documents/MET-G-NLDN-Brochure-B210412EN-E_Low.pdf

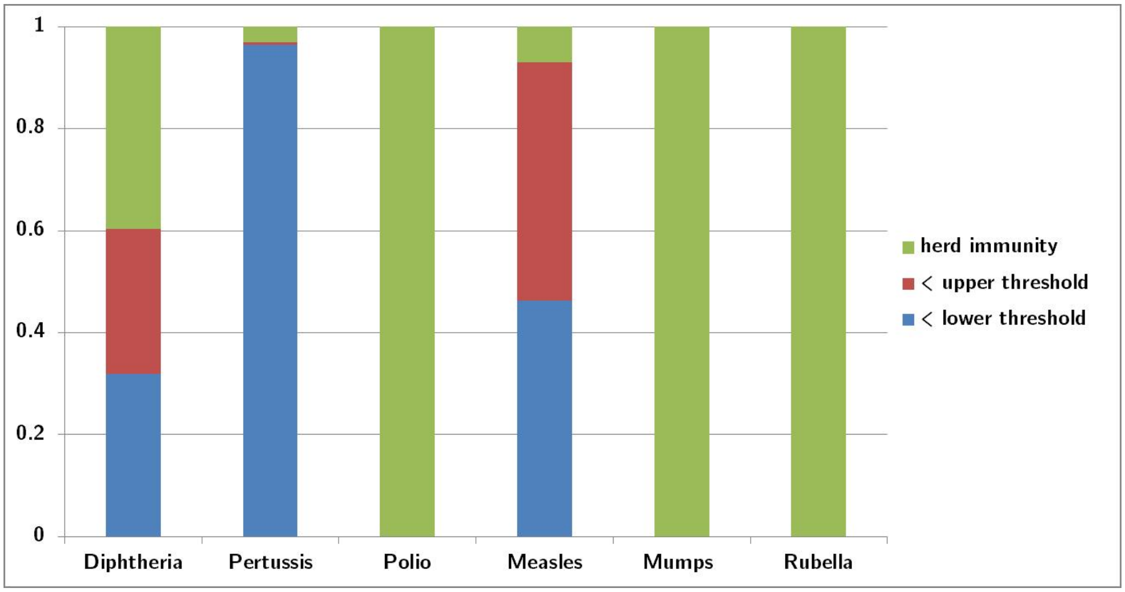

Figure 1: Proportion of counties with herd immunity

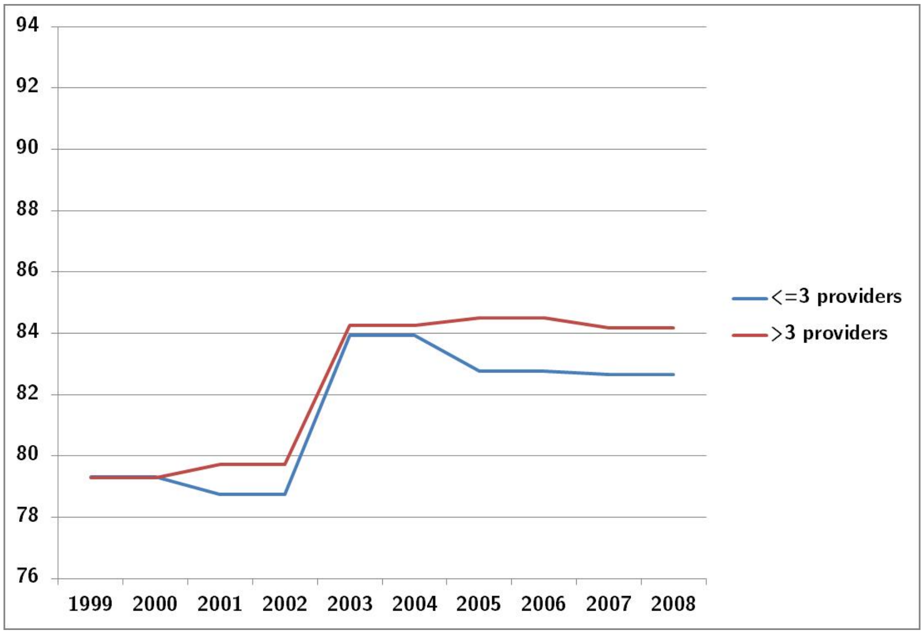

Figure 2: 4:3:1 series, by number of providers

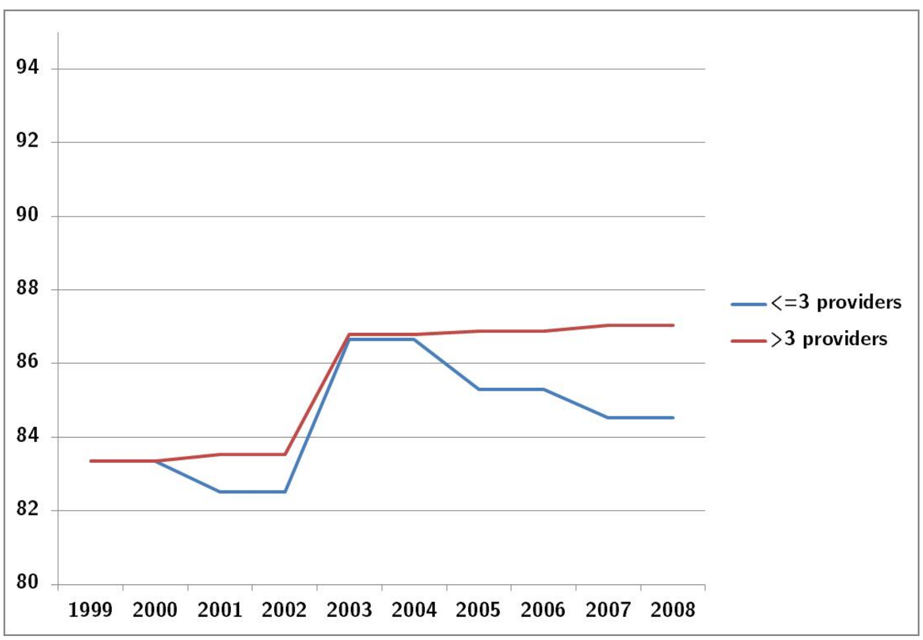

Figure 3: DTaP/DTP series, by number of providers

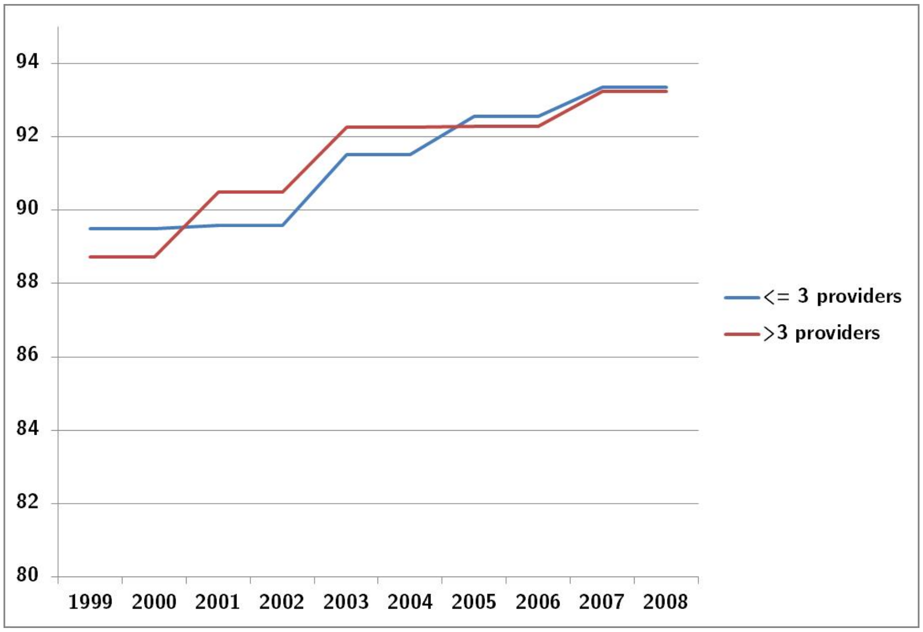

Figure 4: Polio series, by number of providers

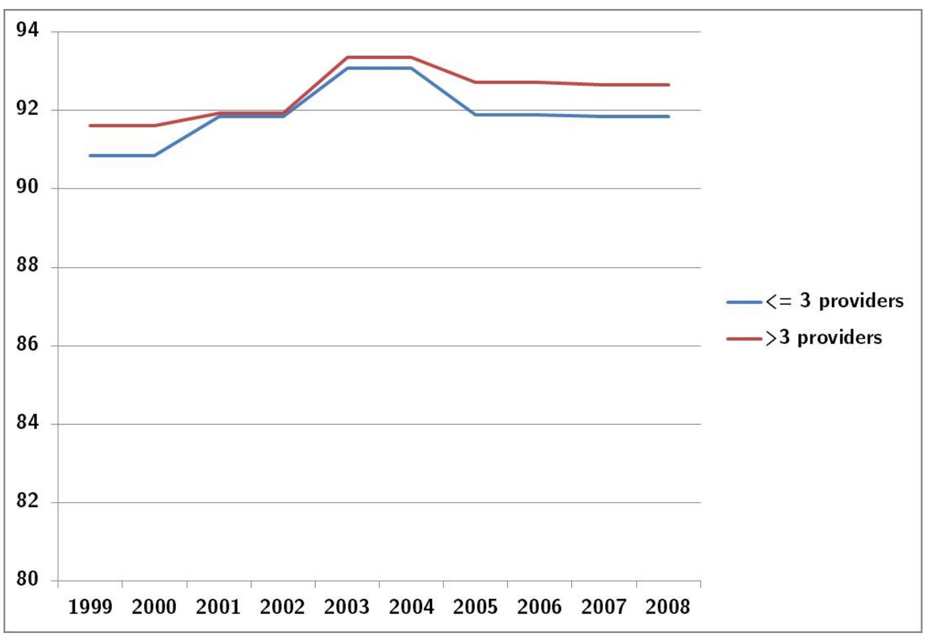

Figure 5: MMR series, by number of providers

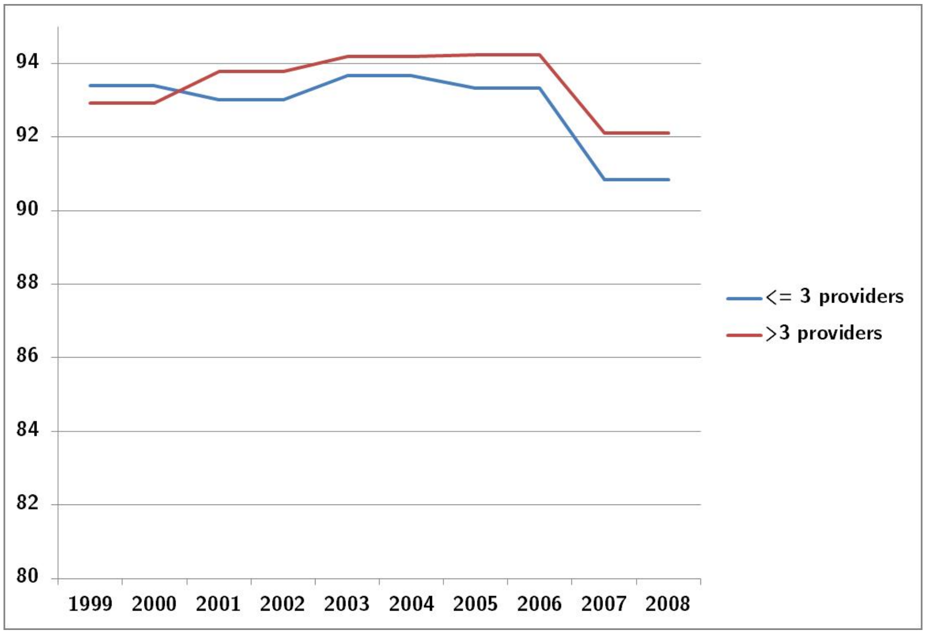

Figure 6: Hib series, by number of providers

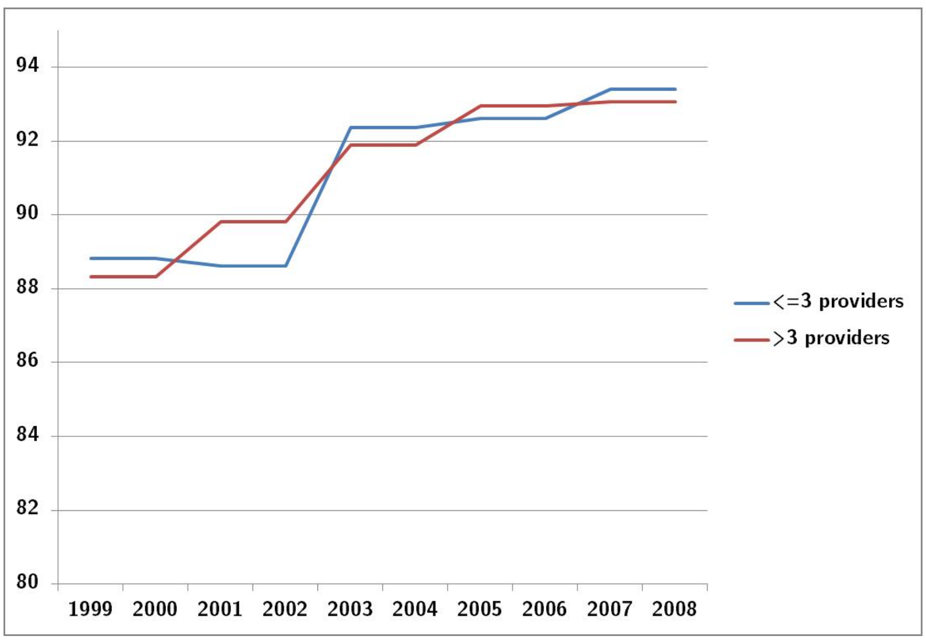

Figure 7: Hepatitis B series, by number of providers

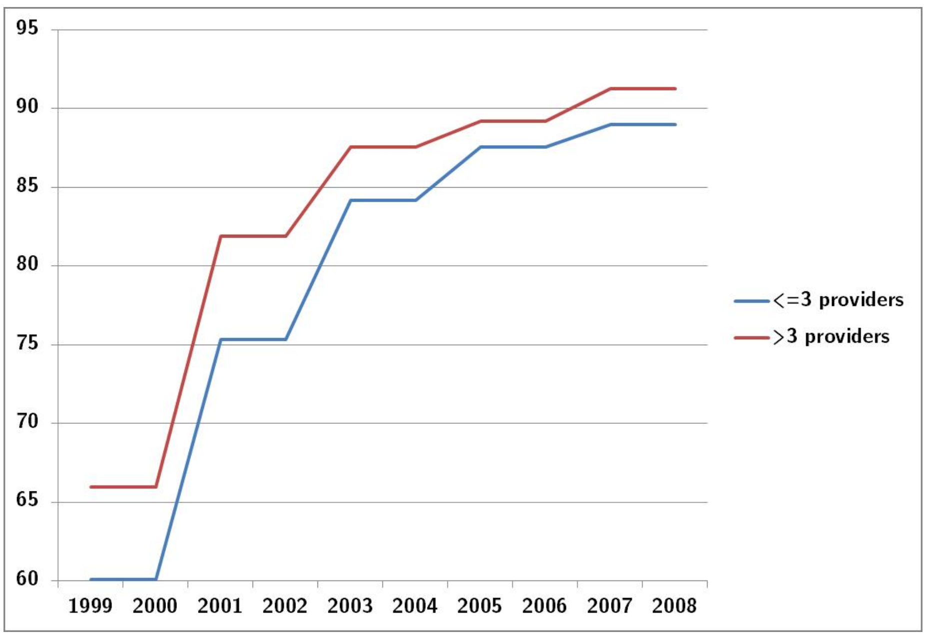

Figure 8: Varicella series, by number of providers

Table 1: Descriptive Statistics

	Mean	Std. Dev.	Min.	Max.	N
Number of HS providers	6.09	4.13	0.01	20.38	2570
Log lightning strikes	7.96	2.03	0	11.85	2570
Log population	12.74	1.12	9.27	16.1	2570
Log pop. density	5.9	1.72	0.68	10.79	2570
Log median HH income	10.76	0.24	10.11	11.58	2570
Log per capita income	10.08	0.24	9.20	11.06	2486
Poverty rate, families with children	11.66	6.56	0.5	43.9	2473
Employment rate	62.83	5.61	41.54	77.38	2562
College degree or higher (%)	29.15	9.17	7	69.40	2570
Share foreign (%)	9.62	8.62	0.58	50.94	2570
Share black (%)	13.79	19.75	0	97.88	2570
Share hispanic (%)	10.41	13.26	0.09	89.60	2570
Information emp. (%)	2.23	1.01	0.52	7.85	2570
Finance emp. (%)	5.65	2.49	1.58	17.26	2570
Professional emp. (%)	7.81	3.46	2.22	24.57	2570
Education emp. (%)	15.95	5.09	7.08	37.12	2570

Table 2: Coverage Rate by Vaccine, 1999-2008

	Min.	Max.	Mean	% Change
4:3:1 series *4 DTaP/DTP, 3 Polio, 1 MMR*	66.1	95.6	81.55	+4.75
DTaP/DTP *4 doses Diphtheria-Tetanus-Pertussis*	66.2	97	84.60	+2.1
Polio *3 doses Polio*	78.1	96.8	91.23	+4.56
MMR *1 dose Measles-Mumps-Rubella*	84.9	96.6	91.99	+1.18
Hib *3 doses Haemophilus influenzae*	79.2	97.3	92.97	-2.38
Hepatitus B *3 doses Hepatitis B*	76.7	96.6	91.06	+5.34
Varicella *1 dose Varicella*	29.2	96.5	79.34	+49.55

Table 3: Herd Immunity Threshold, by Vaccine-Preventable Disease

	${R_0}^a$	${q_c}^b$ (%)
Diphtheria	6-7	83-86
Pertussis	12-17	92-94
Polio	5-7	80-86
Measles	12-18	92-94
Mumps	4-7	75-86
Rubella	6-7	83-86

[a] Reproduction number
[b] Herd immunity threshold, calculated as $1 - \frac{1}{R_0}$
Data: Fine, 1993.

Table 4: Comparison of Constructed Provider Data With Other Internet Measures

	2008	**2008-09**	**2008-11**
Res. HS lines	0.460	0.462	0.434
Res. HS lines (BTOP/BIP Definition)	0.522	0.527	0.501
Total providers	0.800	0.800	0.791
Residential providers	0.616	0.615	0.604
Mobile providers	0.496	0.575	0.662

Notes: Values are correlation of the constructed measure of total number of internet providers per county in June 2008 with the listed variables in December 2008, 2008-2009 (average of three observations), and 2008-2011 (average of six observations).

Table 5: Comparison of FCC-Reported Provider Data With Other Internet Measures

	2008	**2008-09**	**2008-11**
Res. HS lines	0.534	0.536	0.512
Res. HS lines (BTOP/BIP Definition)	0.570	0.577	0.558
Total providers	1	0.991	0.983
Residential providers	0.804	0.615	0.800
Mobile providers	0.491	0.600	0.798

Notes: Values are correlation of the FCC-reported total number of internet providers per county in December 2008 with the listed variables in December 2008, 2008-2009 (average of three observations), and 2008-2011 (average of six observations).

Table 6: Balance on Covariates of the Lightning Instrument

	(1)	(2)	(3)
	Mean- Low	Mean- High	Difference
Population	12.729	12.751	-0.0226
Median HH income	10.803	10.712	0.0906***
Family poverty rate	10.876	12.434	-1.557***
Employment rate	62.937	62.731	0.206
College degree	29.989	28.318	1.670***
Foreign	11.147	8.086	3.061***
Black	12.162	15.411	-3.250***
Hispanic	9.483	11.341	-1.858***
Information emp.	2.222	2.240	-0.0181
Finance emp.	5.574	5.720	-0.146
Professional emp.	7.880	7.737	0.143
Education emp.	16.175	15.717	0.458*
Observations	1284	1286	2570

Table 7: Internet Access and Immunization Status

	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	OLS	IV 1st	IV 2nd	IV 1st	IV 2nd
Number of HS providers	.149**	.123*		.665**		.756**
	(.071)	(.067)		(.338)		(.336)
Lightning strikes			.316***		.294***	
			(.062)		(.047)	
Population		-1.815			.542	-2.056
		(2.292)			(1.654)	(2.709)
Median HH income		-2.102			-8.740***	3.427
		(3.868)			(1.441)	(4.824)
Family poverty rate		-.049			-.038**	-.026
		(.041)			(.018)	(.042)
Employment rate		.123			-.039	.146
		(.101)			(.038)	(.107)
College degree		-.019			.047**	-.053
		(.033)			(.024)	(.041)
Foreign		.106			.066	.061
		(.137)			(.075)	(.144)
Black		-.037			.014	-.047*
		(.026)			(.012)	(.025)
Hispanic		.057			-0.076*	.106*
		(.061)			(.039)	(.064)
Information emp.		-.193			.102	-.261
		(.276)			(.140)	(.291)
Finance emp.		-.298*			.136*	-.392**
		(.175)			(.076)	(.183)
Professional emp.		.398***			.129***	.322**
		(.116)			(.050)	(.134)
Education emp.		.279***			-.008	.283***
		(.073)			(.036)	(.075)
F-Test			26.15		39.47	
R-squared	.457	.499		.417		.442
N	1990	1936	1990	1990	1936	1936

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: The dependent variable in columns (1), (2), (4) and (6) is proportion of children aged 19-35 months that have completed the 4:3:1 vaccine series (4 doses of DTaP/DTP vaccine, 3 doses of polio vaccine, 1 dose of MMR vaccine). The dependent variable in columns (3) and (5) is number of high-speed Internet Service Providers. All specifications include county, state and year fixed effects. Standard errors, clustered by county, are in parentheses.

Table 8: Internet Access and Immunization Status, by Vaccine

	(1)	(2)	(3)	(4)
	OLS	OLS	IV	IV
DTaP/DTP	.066	.036	.249	.367
	(.069)	(.065)	(.322)	(.307)
R-squared	.312	.353	.304	.331
Polio	.127***	.147***	.572***	.732***
	(.042)	(.042)	(.183)	(.224)
R-squared	.548	.566	.480	.456
MMR	.043	.024	-.065	-.089
	(.033)	(.032)	(.147)	(.159)
R-squared	.253	.280	.244	.271
Hib	.098**	.072**	.315	.267
	(.040)	(.035)	(.228)	(.233)
R-squared	.319	.337	.294	.317
Hepatitis B	.135***	.145***	-.064	-.027
	(.049)	(.049)	(.297)	(.341)
R-squared	.589	.592	.580	.585
Varicella	-.383***	-.250**	-1.618	-1.424
	(.131)	(.118)	(1.017)	(1.002)
R-squared	.879	.886	.860	.869
County controls		X		X
N	1990	1936	1990	1936

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: The dependent variable in each panel is proportion of children aged 19-35 months that have completed a specific vaccine series. The independent variable is number of high-speed Internet Service Providers. The county-level covariates are log population, log median household income, poverty rate of families with children under age 18, employment rate, share of population with college degree or higher, share foreign, share black, share hispanic, and employment by industry (information, finance, professional and education). All regressions include county, state and year fixed effects. Standard errors, clustered by county, are in parentheses.

Chapter 2:
Spillovers in Local-Level Conflict

Carrie Huffaker
Università Commmerciale Luigi Bocconi

October 6, 2017

1 Introduction

> Things fall apart; the center cannot hold; Mere anarchy is loosed upon the world, The blood-dimmed tide is loosed, and everywhere The ceremony of innocence is drowned.
>
> from "The Second Coming," W.B. Yeats

Over the last three decades, war has claimed two millions lives in battle, and tens of millions more lives among civilians. Nearly 1000 unique conflicts have been fought since 1989. The vast majority of these conflicts are intrastate, that is civil wars, with frequent intervention by international actors. The Syrian and Iraqi civil wars and rise of Islamic State have caused hundreds of thousands of deaths and displaced millions more. The worst refugee crisis since World War II is still unfolding.[1]

These conflicts have provoked the radical transformation of geographical boundaries, resource allocations, and social identities. The rising tide of terrorist attacks throughout the Middle East and North Africa, spilling into Europe, demonstrates how the consequences of war cannot be contained by well-delineated borders on a map. Indeed, every civil war is at once an international crisis, as refugees, fighters and even supplies of warfare move across space- together with inchoate yet no less powerful forces of desperation, fear and resentment. War is a time of great flux. Wherever it touches, old alliances and old identities are hardened, become *real*, or are supplanted by new and completely unforeseen ones. The social ties binding together a society fall apart.

What does it mean to be a Christian or Muslim in Egypt, or Hutu or Tutsi in East Africa? Such questions provoke far different answers now than they did prior to 2013, when Abdel Fattah al Sisi seized control of Egypt in a military coup, or 2008, when the Second Congo War broke out. Which groups win or lose, and how does the distribution of political power across society shape conflict? Which regions win and lose, and why?

Since the pioneering work of Saleyhan and Gleditsch (2006), which examined the role of refugees in the spread of civil war, researchers have focused new attention on the role of transnational forces in the outbreak of new conflict and intensification of old ones. An important step forward in this field was the introduction of spatial statistics, originally developed in the context of fields like epidemiology, biology and geography. Spatial autocorrelation models explicitly analyze interdependence in the outcomes of connected actors.

The spatial approach shares many similarities with graph theory and network analysis. At the same time, spatial models allow for the incorporation of risk factors shown in conflict literature to be associated with war. In this way, spatial autocorrelation models allow for the joint study of the internal and external processes that contribute to civil conflict, and in particular how conflict in one region influences the outcomes of all the other regions to which it is connected- directly and indirectly. Although not without controversy, the models provide a new and powerful tool for embedding the study of domestic conflicts in an international framework.[2]

[1] These figures are estimated on the basis of data from the Uppsala Conflict Data Program (UCDP), 2016.

[2] See Gibbons and Overman (2010) and Corrado and Fingleton (2011) for two economic critiques of spatial models.

The fundamental component of spatial econometric models is the spatial weight matrix, W. By systematically varying the weight matrix used in the analysis, it is possible to define spatial networks in different ways, and so investigate and contrast different patterns of spatial dependence. The unique power of spatial analysis lies in this ability to explicitly evaluate distinct hypotheses about how conflict spreads.

In this paper, I focus on spillovers of a particular type of conflict: Local-level internal conflict. This encompasses the violence of civil wars, involving government and different rebel movements, but also political conflict that falls outside of officially recognized wars: One-sided violence by governments against civilians, by militias against civilians, and between different groups in society, with no organized state or militia component. This is the full spectrum of violence that societies face today. I also consider conflict of a different type, which is generally nonviolent. Yet carried out on a large scale, these may lead to violence, as has occurred in the Arab world since 2011.

The area of study is the entire African conflict, both north and south of the Sahara, from 1997 to 2015. I show that during this period, there was an important spatial autocorrelation component to the spread of conflict. When a neighboring region experiences greater conflict, so do I; and this spillover effect is even more potent when the two regions are linked by ethnic ties. This holds true even when the impact of geographic proximity is accounted for. Ethnic and sectarian identities are not the only cause, nor even the primary one, of most wars; but they are very often present when bloodshed crosses borders.

I also show that political forces play a role in the diffusion of conflict. In particular, the more distant a region is from the national capital, the more insulated it is from the spread of conflict. Moreover, when a district gains greater access to central power, the entire neighborhood experiences greater political conflict in the form of protests and rioting.

The organization of the paper is as follows. In the next section, I address spatial dependence in outcomes and describe different mechanisms through which it might occur. I also address what sort of districts might be particularly vulnerable to conflict spillovers. In section 3, I introduce the three main spatial models to be estimated. These are the Spatial Autoregressive Model, the Spatial Durbin Model and the Spatial Error Model. In section 4, I describe the data and methodology used in the paper. The presentation of estimation results from all of the different models follows in the next section. Robustness checks are performed in section 6, and conclusions follow in section 7.

2 Theoretical framework

2.1 Spatial dependence in outcomes

The first challenge in analyzing a complex phenomenon like diffusion is to clearly define and identify it. As Gleditsch (2007) writes, "We would like to know not only that conflict in Rwanda increases the risk of conflict in Zaire, but *what it is* about conflict in neighboring states that increases the risk of conflict in another state" (emphasis added). Are the outcomes of Rwanda and the Congo linked because agents travel and communicate across borders? Are actors in one state emulating strategies that they've seen successfully enacted in another? Or are agents in each country behaving in a completely autonomous way, as many past studies of conflict implicitly assume? A clear dichotomy arises between interaction and interdependence, in which case it is appropriate to speak of diffusion; and independent outcomes, in which case it is not.

The clustered adoption of specific practices can occur under three distinct circumstances (Elkins and Simmons, 2005). First, clusters can reflect *coordinated effort* on the part of actors. When different actors engage in strategic interaction, their outcomes are then jointly determined. Many suicide bombings across the Middle East and elsewhere have been carried out by militants loyal to the Islamic State, a clear example of international cooperation towards a common end. The adoption of the euro is a more benign example. Coordinated effort may also arise from coercion. The universal adoption of burqas by women in the lands controlled by Islamic State, and more generally the mass religious conversions that have followed many territorial conquests, are both examples of clustering-by-coercion. In all cases, both actors are actively involved in transmitting contagion.

Second, clustered activity may arise due to *uncoordinated interdependence*: One transition- for example, the outbreak of war in one region- increases the probability that war will start in a second region, through *no conscious intention* on the part of the initiators. The two principal channels of this uncoordinated interdependence are *adaptation* and *learning* (Ibid). Groups adapt to changed circumstances when the initial adoption of a practice increases their own payoff from adopting it; they learn when the initial adoption imparts new information about the cost-benefit trade-off of such an action. When adaptation and learning occur, outcomes are correlated even when actors do not directly interact with each other. The diffusion process tends to be one-sided: the destination actor responds to a new situation which the actor at the origin has created.

Coordinated effort and uncoordinated interdependence both represent theoretically interesting situations of diffusion. In the first situation, the origin and destination actively collude; both actors are directly involved. In the uncoordinated case, the origin may not even be aware that contagion has occurred. The key feature that separates *absence of diffusion* from either scenario is that of *independence*. Groups with similar economic, political and cultural attributes may respond to *like shocks* in a *similar way*. Many of the risk factors associated with conflict, such as poverty and ethnic fractionalization, are clustered in certain geographic regions; moreover, such regions are vulnerable to the same sorts of shocks (Fowler, 2011). In such cases, behavior appears to demonstrate dependence, yet there is no causal link in the spread of a behavior from one observation to another. Each region operates autonomously.

Evidence of diffusion may arise for atheoretical reasons. Spatial processes may operate on a larger scale than the single unit of observation. Administrative boundaries such as states and provinces often fail to correspond to the neighborhoods that give rise to the variables that we can observe. Moreover, latent, unobservable or unmeasurable features that explain variation in the outcome might be spatially autocorrelated. In both cases, observed relationships may be statistically significant, even in the absence of an underlying theoretical explanation (Buhaug and Gleditsch, 2008; LeSage, 2004).

The identification of diffusion requires us to distinguish between like regions responding to like circumstances, statistical anomalies, and true relationships of spatial dependence, which are characterized by interaction and interdependence. It can be empirically difficult to distinguish between the cases of coordinated and uncoordinated interdependence. We usually lack the information needed to establish clear and definitive ties between actors, or to impute intentions to their behavior. Few states would admit to purposely destabilizing inimical neighboring regimes by aiding rebel movements, for example. In the coordinated case, one agent [3]

In this paper I define diffusion as *spatial dependence in the outcomes of two connected districts.* It is a process linking together an initial stimulus in one region with behavior in another, through some sort of medium or conduit and through the choices of specific social agents (Solingen, 2012). Which mechanisms favor transmission, what characteristics render regions vulnerable, and how regions are linked together therefore jointly determine when spillovers will occur, and when they will not.

2.2 Spillover mechanisms

The mediums by which spillovers occur are of two types: soft and hard. *Soft channels* produce changes in information and incentives, which in turn spark adaptation and learning. When actors observe a successful (failed) revolt abroad, they use this new information to update their priors about the probability of success and thus are more (less) prone to start a revolt in their own country. This is known as the demonstration effect, an important channel of diffusion. The global games framework- in which the observed signals of other players impart information about their beliefs and likely actions- is closely related. A group may be pushed into dangerous and illegitimate behavior based upon the belief that others are engaging in it (Morris and Shin, 2001). There is strength in numbers: The greater the number of rioters or protesters, the lower the probability that any individual participant will suffer punishment (Hardin, 1995). Yet such beliefs can be misleading, leading to the buildup of waves of actions that are not beneficial ex-post and not reflective of the true state of the world (a strong and resistant regime, for example, as opposed to a weak and vulnerable one). The failed uprisings and ensuing civil wars in Yemen, Syria and Libya are possible cases of the demonstration effect leading to suboptimal outcomes. For even as opposition movements observe events abroad, a second diffusion process is occurring simultaneously: Authoritarian learning on the part of rulers (Lynch, 2014; Brownlee et al., 2015; Danneman and Ritter, 2014). Anticipating the shifting information available to their own opposition movements, regimes may engage in preemptive repression in order to avoid the same undesirable outcome of coup d'etat. Conflict can therefore spread to new destinations through the behavior of either side, the rulers and their allies or opposing forces.

The demonstration and global games effects require no interaction between origin and destination agents: The destination must merely be aware of and able to observe the origin (the rationale behind severe restrictions on internet access in North Korea, see Zeller (2006). Yet a second soft channel- that of communication- requires the active participation of both origin and destination. Effective communication is essential to overcoming one central strategic challenge to collective action, that of coordinating potential participants. The expansion of

[3] Yet evidence of such aggressive strategies is widespread, especially in Africa. In 2012, the Hague's Special Court for Sierra Leone convicted former president Charles Taylor of Liberia of supplying and encouraging rebel atrocities in Sierra Leone, the most egregious example of this strategy. Throughout the 1980's, South Africa's apartheid regime worked to overthrow newly elected Robert Mugabe in Zimbabwe, contributing to his paranoid mistrust of Europeans and whites (Meredith, 2007).

social networks and mobile coverage has enabled protest movements to grow in Russia and Africa (Enikolopov *et al.*, 2016; Manacorda and Tesei, 2016); cellular communication has also aided counterinsurgency efforts in Iraq, by providing a safe and secure means for civilians to provide anonymous tips about insurgent violence (Shapiro and Weidmann, 2015). Communication between actors in different locations allows them to share information and coordinate their efforts on a larger scale, leading to interdependence between outcomes in the two locations. Such links therefore represent a second indirect channel of diffusion. Weidmann (2015) finds that communication networks aid the international spread of ethnic violence.[4]

While soft channels are often unobservable, *hard channels* of diffusion are the highly visible fruits of warfare: The steady flow of fighters, arms and other materiel across administrative borders; the flood of humanity escaping out of conflict zones and into refugee camps. While the first type of flow is difficult to measure, a number of studies have found a significant link between refugees and diffusion (Salehyan and Gleditsch, 2006; Salehyan, 2007; Buhaug and Gleditsch, 2008; Bohnet, 2012; Ruegger, 2013). Studies of migration due to environmental stress identify several important mechanisms: Migrants upset the existing balance of ethnic groups in the house country, heightening intergroup tensions; they also exacerbate competition for resources such as jobs and housing (Reuveny, 2007; Warnecke et al., 2010). Conflict refugees present additional challenges: They abound in grievances and have a low opportunity cost of fighting, and refugee camps often provide sanctuary to foreign rebels (increasing risk in country of origin as well). [5]

Hard channels are particularly important for the diffusion of violent conflict. Because transportation costs are increasing in distance, these hard channels are more likely to operate at close geographic proximity, and are less important for diffusion across longer distances.

2.3 Who is at risk, and why?

Social agents are responsible for propagating diffusion. Yet not all types of agent are likely to do so. The characteristics of origin localities and of potential destinations, and the links that bind them together, are crucial in determining which districts are vulnerable and where conflict is likely to spring up next. Much of the foundational work in conflict spillovers modelled probability of outbreak in a target country as a function of own characteristics and those of neighboring countries (potential originators of conflict). The distribution of ethnic groups making up the population, whether these groups had kin abroad, political institutions, and type of conflict have all been shown to be of first-order importance.

States exhibiting polarization- when a large ethnic minority confronts an ethnic majority- are at greater risk of contagion (Forsberg, 2008), although there is also support for fractionalization- a large number of small groups- as a risk factor (Gleditsch, 2007). Similarly, states are vulnerable when their domestic ethnic groups transcend national borders. The dangerous influence of transnational ethnic kin (TEK) groups is a strong

[4] Networks characterized by more direct links (trade and migration) are not significant, although this is at odds with findings elsewhere in the literature.

[5] Cross-border rebel sanctuaries have played a role on both sides of the Rwandan civil war. The Rwandan Patriotic Front, composed of Tutsi refugees, invaded Rwanda from Uganda in 1990 and 1994; refugee camps in Kivu (eastern Congo) hosted many of the Hutu paramilitaries and government officials who carried out the 1994 genocide, and their presence contributed to the outbreak of war in Congo in 1998. See Gourevitch, 1998.

result found across the literature. Such heightened risk of conflict exists even in the absence of contagion; the very existence of TEK threatens domestic stability, whether or not these TEK are engaged in conflict in their own states. The threat in increasing in the size TEK groups, but only up to a certain point; very large groups (such as the Russians or Chinese) are able to suppress rebellion among their kin in neighboring countries (Cederman *et al.*, 2009; Cederman *et al.*, 2013).

Access to power also matters. Dyadic and triadic analyses show that when a domestic ethnic minority gains access to central power, the probability of conflict decreases; downgrading or complete alienation power instead increases risk. Discriminated and separatist minorities pose the greatest threat of all. Moreover, TEK excluded from power in their own country are more dangerous for the domestic arena, perhaps because they have less of an incentive to maintain regional stability (*Ibid*). Similarly, separatist conflicts abroad- especially those involving TEK- are more destabilizing than are conflicts over central power (Buhaug and Gleditsch, 2008; Brathwaite, 2014). Analyzing Cold War era international disputes, Woodwell (2004) finds that shared ethnic majorities create conflict over group leadership and ideology, while ethnic groups that are a minority in one state and a majority are seen as a threat to state sovereignty and territorial integrity.

Domestic economic and political institutions play an important role in blocking or favoring diffusion of unrest and conflict. States that are more integrated into the regional economy are less at risk of spillovers. Greater state capacity- as proxied by the ability to control and extract resources- can also serve to suppress them (Braithwaite, 2010). This capacity can be translated into coercion or appeasement of the local population, two alternate routes of diffusing potential spillovers. The relationship between political openness and risk is not monotonic: Autocracies with elected legislatures experience less conflict, until conflict breaks out abroad, at which point they are in greater danger (Maves and Braithwaite, 2012). The politically stagnant but oil-rich Gulf Monarchies were able to head off Arab Spring by "buying off" the opposition (through rich concessions), while less wealthy mixed regimes that attempted to introduce some measure of democracy proved far more vulnerable (Brownlee, 2015). In Syria, violent opposition failed to spread to those districts that had been targeted for selective redistribution by the Assad regime (De Juan and Bank, 2015). Finally, recent studies of local-level conflict dynamics find a negative relationship between spending on development and reconstruction and insurgent violence in Afghanistan, Iraq and elsewhere (Berman et al., 2011; Berman and Matanock, 2015).

As the above discussion suggests, many of the same risk factors for independent outbreaks of conflict- including poverty, ethnic divisiveness and instability- also place a country at risk of contagion. Yet controlling for domestic factors alone cannot fully capture this risk. As Gleditsch (2007) shows, "the likelihood of civil war in an extremely unfavorable region (is) several hundred percent higher than the risk of conflict in very favorable neighborhoods." Such assertions are clearly sensitive to one's definition of neighborhood. Having discussed different channels of diffusion and social actors, I now address the central issue of diffusion analysis: How to define connectivity. Past studies have adopted the criteria of geographic contiguity, distance, and finally relationships based upon other characteristics. Danneman and Ritter (2013) analyze contagion at the national level, using spatial weight matrices based upon distance, religion and culture, and political similarities.

While political institutions and policies are endogenous to conflict processes, ethnicity, language and religion are slow-moving variables that provide a strong contrast to contiguity-based measures of similarity. Moreover, such variables have economic and theoretical meaning. Spolaore and Wacziarg (2015) show that more closely related populations tend to go to war more frequently, due to similar preferences over rival goods and the lower cost of ruling over similar groups (creating an "economies-of-scale" incentive for territorial conquest that Alesina and Spolaore (1997) also address). These results are robust to redefining *relatedness* in religious, linguistic and cultural terms, which Spolaore and Wacziarg demonstrate in a 2016 paper to be both closely related and easily summarized by the genetic distance between populations. Ethnicity is therefore an objective characteristic with clear biological and historical foundations.[6]

This attribute shapes behavior because individuals derive utility not only from their own actions, but those of other group members. They enjoy benefits and bear costs that accrue to them personally, yet are at the same time sensitive to the power, status and reputation of the group as a whole (Akerlof, 2000). When group membership becomes part of the self, "events that harm or favor an in-group by definition harm or favor the self, and the self might thus experience emotions on behalf of the in-group" (van Stekelenburg and Klandermans, 2013). Not all group memberships have salience for the individual all the time. A distinction can thus be established between collective social identities that are still latent, and therefore separate from the self, and those that shape individual behavior. Sambanis and Shayo (2013) present a model of endogenous social identification in which individuals choose whether to align themselves with a national or subnational identity, as a function of group status and the permeability of group boundaries. This process of individual and social alignment- the *politicization* of identity- is a crucial step in transforming group membership into collective action: The riots, protests and violence under analysis here (Simon and Klandermans, 2001). Yet most models of diffusion have failed to allow for the degree to which self-identity and collective social identities coincide or contrast.

Politicized identities are a source of conflict when power structures are unclear or unstable (*Ibid*). Rationalist explanations for war posit that peace will break down only in situations of incomplete or asymmetric information. States uncertain about their opponent's ability or willingness to use force are compelled to use force themselves, as the bargaining that could fend off war becomes difficult if not impossible (Gartzke, 1999). As discussed above, refugee flows create uncertainty by upsetting a region's existing demographic balance. Shocks concentrated abroad, but involving ethnic kin- either their arrival to power in a neighboring state, or their engagement in a war- can be similarly disruptive. The politicization of identity means that such shocks can alter information and incentives, activating the soft channels of diffusion. This can occur even when the destination is not directly impacted by the original conflict; even when two states are geographically distant.

When transnational kin become embroiled in conflict, domestic ethnic groups may perceive a change to their status quo that encourages them to act. Rival groups may exercise preemptive force to stave off the future strengthening of their opponent (Danneman and Ritter, 2013). Each sort of shock creates uncertainty about the true strength of a group, a problem of asymmetric information that can lead to bargaining failure and

[6] Hirshleifer (1998) argues that kinship and group membership arose through evolution as a "returns to scale" in contending for ultimate biological goods such as access to food and reproduction.

the outbreak of conflict (Jackson and Morelli, 2009).

Yet clearly not all- nor even most- conflict is identity-driven. The importance of identity politics can hold true even when local-level conflicts, such as those analyzed here, have very little to do with macro-level cleavages. As Kalyvas (2003) argues, civil wars are the joint production of the strategic actions of central actors and the opportunistic actions of local ones. While the latter recruit and mobilize supporters, central actors inject extra muscle into local conflicts, creating a "concatenation of multiple, often disparate local cleavages, more or less loosely arranged around the master cleavage." Nor must local identities fully align with collective ones in order for these actors to benefit from group success (Hardin, 1995). Finally, targeted violence against civilians can render nonparticipation in a conflict so costly that civilians have little choice but to participate: An alternative solution to the collective action problem (Kalyvas and Kocher, 2007). In this way, violence contributes to the hardening of collective identities and master cleavages, as civilians seek refuge in those groups less likely to do them harm.

As the existing body of literature demonstrates, conflict of all kinds spills across space. Such diffusion effects remains significant even when the spatial clustering of risk factors is accounted for. Mechanisms may be soft or hard, and different types of mechanisms are likely to favor different types of conflict. Originators of conflict, as well as targets, may be linked by geography or by shared ethnic communities. Whether these communities have access to power in either state is an important consideration. Finally, ethnicity represents an important motor of collective action, even when local actors are pursuing their own self-interest.

The existing literature on diffusion falls short in a number of ways. National-level studies address both ethnic identities and access to power, but on such a large scale as to confound the true dynamics and driving force of conflict: The behavior of local actors, and how this varies across time. Addressing diffusion on a national level also neglects the rich regional heterogeneity that drives most conflict in Africa. I therefore focus upon regional-level dynamics, adopting first-level administrative boundaries (states, provinces and districts) as the basic units of analysis.

Past studies of local spillovers include Harari and La Ferrara (2014), who focus upon how rainfall shocks generate waves of conflict at an extremely disaggregated (cell) level of analysis; and Konig *et al.* (2015), who analyze strategic dependencies in fighting networks during the Second Congo War. My work differs from theirs in directly evaluating the impact of political power and ethnic identity upon the process of diffusion. I integrate the theoretical ambitions of a national-level approach and the sharp focus of a local-level one into a single framework, and apply this framework on a very large scale: All of Africa from 1997-2015.

As far as I am aware, this is one of the first attempts to undertake a comprehensive view of how ethnicity and political power influence the diffusion of conflict from region to region. I focus upon six basic hypotheses that arise out of transnational literature, yet have never been tested on regional-level data. The first two contrast geographic and ethnic ties in their ability to serve as mediums for diffusion. The third addresses different types of conflict. The final two hypotheses address the role of political power and peripherality.

(H1) Local-level conflict processes display geographical contagion, even once the spatial clustering of risk factors is accounted for.

(H2) Ethnicity is an important channel for contagion, beyond the effect of geographic contiguity.

(H3) Diffusion of non-violent conflict, like protests and demonstrations, is likely to be driven by different determinants than diffusion of violent conflict. In particular, geographic contiguity- which favors hard channels of diffusion- is likely to be more important for violent conflict.

(H4a) Districts that are isolated and peripheral will be more vulnerable to spillovers than will districts with access to political power.

(H4b) Conflict shocks originating from powerful districts, versus marginalized ones, will provoke a more powerful response.

3 Modelling contagion

When spatial dependence is present, OLS estimates are biased and inconsistent, as well as inefficient.[7]

To address this issue, models of spatial autocorrelation were first introduced into the sociological and geographical literatures in the 1950's. Such models represent an extension into space of the older autoregressive models used to analyze time series data. In the case of time, dependence flows in a single, forward direction; but more complicated structures hold in spatial models. Dependence can flow in two directions at once through space, it is often asymmetric, and it triggers feedback loops.

As in time series analysis, the first step in spatial modelling is to test whether such an autoregressive approach is even appropriate, by analyzing residuals from an ordinary regression for residual spatial correlation. Moran's I and Geary's C are the two primary tests used to evaluate global and local spatial autocorrelation in data (Moran, 1950; Geary, 1954). When a data generating process exhibits significant autocorrelation, the fundamental tool of spatial econometrics to model this dependence is the spatial weight matrix, W. Like the adjacency matrix in graph theory, W represents links between the elements of a given system or network. The spatial weight matrix may be binary or row-stochastic, in which case rows are normalized to sum to one. The characteristic roots of the matrix must be known- it is non-parametric- and elements are not allowed to be neighbors of themselves (diagonal elements equal zero). Weight matrices may be asymmetric (one-way, directed relationships) or symmetric (undirected relationships that permit possible feedback loops).

The key feature of the W matrix is its flexibility. By defining and redefining the matrix according to different criteria, it is possible to compare and contrast competing spatial hypotheses in an analytically rigorous way. As Corrado and Fingleton (2011) argue, "where spatial econometrics takes a lead is in its ability to identify and test theory relating to explicit spatial dependence mechanisms, as embodied in the parameterization of a W matrix." Neighborhoods defined by contiguity or by some limiting distance from the target (1 km, 750 km, etc.) enable us to test *direct* mechanisms. These are likely to coincide with the hard channels of diffusion discussed above. Refugees flee to to the nearest international border, in general. When sea, air and road

[7] When spatial lags in the dependent or in the regressors are significant (the SAR and Spatial Durbin models, discussed below), OLS estimates are biased; while in the SEM case of spatially correlated errors, the OLS estimator is merely inefficient (Elhorst, 2003).

networks are relatively underdeveloped- as in Africa- the transport of men and arms becomes more costly. For such channels, distance and nearness matter.

We can test *indirect* mechanisms by defining W in more imaginative ways. Religious neighbors share a common faith; ethnic neighbors, common culture and practices; political neighbors, common policy preferences. If these relationships remain significant even when contiguity and spatial distance been accounted for, then we have direct evidence of their validity in determining joint outcomes across space. In the context of conflict, we can identify a region's potential allies (like, linked regions) and enemies (unlinked regions). Although distance is distinctly important for both soft and hard channels of diffusion, social, political and cultural allies are likely to be especially susceptible to the soft channels. Collective identities provide a potential basis for group membership. There is likely to be greater communication between groups that share similar languages and history; on the basis of their shared identity, they may observe and emulate each other in a special way.

Once the W matrix has been defined, there are three main ways of modelling spatial autocorrelation. Each one serves a distinct purpose. By introducing lags of the dependent into the set of regressors, we can establish whether spatial diffusion of the outcome is actually occurring. Spatially lagged disturbances control for latent variables that are correlated across space and influence the dependent. Finally, lagged covariates allow for dependence in the outcome of any single region upon not only its own characteristics, but also those of its neighbors. Together, the complete framework of the three models provides three different spatial parameters (λ, γ and ρ), each governing the effect of a distinct type of spillover.

The most basic version of the spatial model posits first-degree autocorrelation in the dependent variable, in correspondence to the temporal AR(1) process. This is the Spatial Autoregressive Model, or SAR, sometimes referred to as the Spatial Lag Model. Given a dataset of N observations and an NxN spatial matrix, W, the SAR model is given by:

$$(I_N - \lambda W)\, y = X\beta + \varepsilon$$

$$y = \lambda W y + X\beta + \varepsilon$$

In reduced form:

$$y{=}(I_N - \lambda W)^{\,-1} X\beta + (I_N - \lambda W)^{\,-1}\varepsilon$$

where I_N is the NxN identity matrix, λ is the autoregressive parameter, and y is the vector of dependents. The dependent is a function of both observable characteristics $(X\beta)$ and unobservables (ε).

By lagging the dependent variable, the SAR model specifies spatial effects as a diffusion process. The model captures both strategic interaction and uncoordinated interdependence. A significant λparameter provides basic evidence that outcomes of neighboring units are jointly determined, and contagion is actually occurring (Sparks, 2013a; Elhorst, 2013). Rebellion in one region might cause neighboring regions to update their probabilities of success, create a new environment that neighbors must adapt to, or spark destabilizing flows of refugees or arms. All of these impacts would be captured by the SAR parameter. As λ rises, spatial effects grow more persistent, approaching a limiting "spatial random walk" scenario in which a shock to the outcome of a single district propagates through the entire system with undying force, and the entire neighborhood becomes “infected” by change.

The Spatial Durbin Model (SDM) introduces the observable characteristics of neighboring districts into the information set. Actors do not behave in isolation, whether they are individuals, districts or entire nations. In undertaking decisions, actors will always account for the larger strategic setting in which they are embedded. Unlike the SAR model, which implicitly assumes that only shocks to neighboring dependents matter, or traditional models that account only for “own” attributes, the Durbin model allows for the attributes of others to directly enter into an individual region’s decision-making process. Spatial Durbin estimates are of particular value because they result unbiased even when the true data generating process is a SAR or SEM.

The spatial Durbin model is given by:

$y = y + X\beta + WX\gamma + u$

where WX is the vector of spatially lagged covariates and γ is the Durbin parameter vector. The coefficient γ_r measures the average impact of a change in variable x_{ir} upon the outcome of region j, $i \neq j$. This parameter reflects indirect (neighborhood) effects, while traditional regression coefficients capture direct effects (LeSage, 2008; Glass et al., 2012). However, the situation is complicated by the fact that the indirect effects of any type of shock will eventually lead back to the district of origin itself. In evaluating the total impact (direct + indirect) of a variable, it is necessary to account for the entire cycle of spillovers and feedback effects that an initial change can generate.

Suppose a politician from region A becomes president. Suddenly, region A has much greater power and influence on the national level. Levels of violence may fall in A, as residents reap the benefits of a shift in the geographic distribution of selective redistribution in favor of their own region. As residents travel outside the region to buy and sell goods, or send remittances, this new wealth spreads to other regions, potentially lowering violence there, as well. At the same time, region B has lost access to the privileges of executive power. As B suffers from lower levels of patronage, violence may rise there and in neighboring regions, which are also negatively affected. If A and B are neighbors, the situation becomes yet more complicated. In each case, feedback effects are generated, which sometimes operate in a different direction from the initial, direct impact of the change in x_{ir}. The average total impact measure reflects the long-run equilibrium effect of such a change upon the entire network. Evaluation of the direct, indirect, and long-run impact of shocks to the dependent is also possible in the SAR model.

In this paper, I estimate both spatial and non-spatial models, in order to understand how the intensity of conflict changes. I use the spatial Durbin model estimates to distinguish between direct and indirect spillovers. Finally, as a robustness check I estimate the SAR in combination with a third model of spatial dependence: The Spatial Error Model (SEM). In the SEM, the error term is assumed to be spatially correlated:

$y = X\beta + u$

$u = \rho W u + \varepsilon$

$y = X\beta + (I_N - \rho W)^{-1}\varepsilon$

The error term has two components, the spatially autocorrelated u and the random, homoskedastic ε. The spatial error parameterρ measures the strength of spillovers among the model residuals of neighboring observations (Sparks, 2013b).

The SEM framework allows us to control for latent variables that are unobservable or difficult to quantify, and that might be correlated with conflict. The basic units of analysis- the administration 1 districts- are of different shapes and sizes, and their borders are possibly endogenous. The same holds true for the homelands of different ethnic groups. Moreover, the spatial process under examination may not coincide with any of these borders. Finally, unobservable idiosyncratic shocks might follow a spatial diffusion process (Elhorst, 2013; Millo, 2014). All of these possibilities give rise to spatial correlation that is distinct from the dependence in outcomes described by a SAR process. Inclusion of the SEM component allows us to distinguish between the effects of th two types of spatial dependence.

By combining the SAR and SEM models, we can analyze whether spatial diffusion in outcomes (as represented by the SAR parameter λ) persists, even after controlling for the influence of spatially autocorrelated unobservables (represented by the SEM parameter ρ). If violence breaks out first in region A and then in neighboring region B, it may appear that B is emulating A, a SAR effect. However, a shock such as plant disease, extreme weather, or a rise in brigandry- all of which would be unobservable to the outside observer- might have first impacted A and then B, threatening livelihoods and lowering the opportunity cost of violence. In introducing the spatially correlated error term, we discern whether true spatial diffusion is at work: Thus avoiding the trap of similar localities responding to similar circumstances.

This type of combined analysis is commonly known as the SAC model. Given a panel dataset of N observations and T time periods, the SAC model with k covariates is given by:

$$y = \lambda \left(I_T \otimes W \right) y + X\beta + u$$

where y is an NTx1 matrix of observed dependents, I_T is the TxT identity matrix, X is the NTxk matrix of covariates and u is the NTx1 vector of spatially autocorrelated residuals. As before, $\lambda$is the SAR parameter and β is the kx1 coefficient vector.

The model with individual effects has a two-part disturbance vector:

$$u = \left(\iota_T \otimes \mu \right) + \varepsilon$$

$$\varepsilon = \rho \left(I_T \otimes W \right) \varepsilon + \nu$$

where ι_T is a Tx1 vector of ones and μ is the Nx1 vector of time-invariant individual effects. As before, ρ is the SEM parameter. The vector of spatially autocorrelated idiosyncratic errors is given by ε, while $\nu_{it} \sim IID\left(0, \sigma_\nu^2\right)$ by hypothesis. The distribution of individual effects μ depends on whether we allow these to be correlated with the explanatory variables (fixed effects), or adopt the simplifying assumption of independence (random effects).

4 Data and methodology

Table 1 presents district-year descriptive statistics for the main dependent and independent variables, described below. Over 18 years, 764 districts were observed, for a total of 13,752 district-year observations.

4.1 Dependent variables

The outcome of interest is event occurrences in Africa from 1997-2015, taken from the Armed Conflict Location and Event Data Project (ACLED). The information about conflict events is gathered from a wide variety of international news sources. The type of event and actors involved are all recorded, as are the precise geographic coordinates of the locality where the event occurred.

Based upon the ACLED dataset, three dependent variables were created. Each measures the number of occurrences of a specific type of conflict event. The third hypothesis is that violent and nonviolent conflict are driven by different types of shocks. By disaggregating conflict occurrences in this way, we can test whether different mechanisms are indeed at play, and whether one type of conflict is more easily transmitted.

By utilizing counts as the outcome, rather than a simple dummy for conflict onset- as most past studies have done- we are focusing on the *prevalence* of conflict in the population of African districts, rather than *incidence* or onset. Such an approach allows us to exploit variation in the intensity of conflict. A count variable allows us to distinguish, for example, between a few contained instances of unrest, and a full-scale civil war that generates numerous casualties and massive refugee flows. Moreover, districts that have experienced conflict in recent years- and are thus already counted as "infected" in the incidence sense- may undergo drastic changes in terms of level and type of conflict, due to spillovers. A prevalence approach allows us to exploit to the fullest extent all of the information available in our dataset.

The first dependent, n_events_{it}, is the total number of conflict events (of any type) that occurred in district i in year t. The second dependent, $riots_{it}$, measures riots and protests, which tend to be nonviolent events in the ACLED dataset.[8] The indicator $violent_{it}$ measures the number of violent conflict events: Battles, remote violence (such as drone strikes), and attacks on civilians all fall into this category.[9]

4.2 Weight matrices

The first hypothesis is that geographically contiguous regions are subject to the diffusion of conflict events; the second, that regions related by ethnicity experience diffusion. I propose to test hypotheses (1) and (2) through the construction of for different spatial weight matrices, G , E, G_R and E_R. The first matrix contains information on *geographic contiguity*: $w_{i,j} = 1$ if districts i and j are neighbors, and 0 if they are

[8]Riots and protests resulted in 0.32 fatalities, on average, with a maximum value of 631 (rioting between Christians and Muslims in Nigeria in 2008). Eight fatalities occurred per violent conflict event, on average, with a maximum value of 25,000: in 1997, the Alliance of Democratic Forces for the Liberation of Congo-Zaire massacred Hutu refugees from Rwanda.

[9]Specifically, the five types of conflict events included in the *violent* category were "Violence against civilians," "Battle-No change of territory," "Battle-Government regains territory," "Battle-Non-state actor overtakes territory" and "Remote violence."

not. The second matrix defines neighborhoods on the basis of *ethnic contiguity*: $w_{i,j} = 1$ if the same ethnic group lives in both districts, and 0 if there are no shared ethnic groups.

The final two matrices isolate the effects of geography and ethnicity by purging relationships based upon the other variable. The *residual geographic contiguity* matrix defines a neighborhood of regions sharing geographic but not ethnic ties. This matrix is derived as $G_R = G - E$. The *residual ethnicity matrix* defines the neighborhood of regions sharing ethnic but not geographic ties, and is derived as $E_R = E - G$.

These four weight matrices can also be used to test hypothesis (3), that due to the higher cost of transporting fighters and resources across long distances, geographical contiguity should be more important to the diffusion of violent conflict. If this hypothesis holds true, than the estimated SAR parameters for the violence model should be larger and more significant when neighborhoods are defined by geography rather than by the less clearly-defined force of shared ethnicity.

Figure 1 maps out the neighborhoods defined by each of the weight matrices. The ethnic contiguity matrix (in grey) is a slightly sparser version of the geographic contiguity matrix (in green), while the residual contiguity matrix (in purple) is sparser still. The residual ethnicity matrix (not pictured) largely resembles the ethnicity one.

The weight matrices are defined based upon the 768 administrative areas of Africa contained in the Global Administrative Areas (GADM) database. These are the basic units of observation. They are the largest sub-national districts in each country, corresponding to states, provinces and governorates. Both their size and shape vary widely from country to country. Burundi, one of the smallest countries in Africa, has 17 such regions; while the DRC, one of the largest, has only nine. In order to account for this heterogeneity, I employ the Spatial Error Model with district-level fixed effects.

Four of the 768 regions have no direct neighbors, and are excluded from the dataset. The geographic contiguity matrix G therefore has dimensions of 764x764. However, spillovers may occur between these regions because they tend to share a large number of ethnic groups- not because they are near in the physical sense. To isolate the role of geography, the residual contiguity matrix G_R excludes 172 more regions to attain dimensions of 592x592. The SAR, Durbin models and employing the G and G_R matrices provide the evidence we need to test the first hypothesis.

The ethnic contiguity matrix, E, is based on the Geo-referencing Ethnic Power Relations (GeoEPR) dataset. The GeoEPR provides polygons that describe the location of settlement areas of politically relevant ethnic groups across the world from 1946 to 2013. International borders play an important role in the composition of the GeoEPR dataset. For example, Hutus in Rwanda, Burundi and the DRC are counted as three distinct ethnic groups. In order to distinguish common ethnic identities that cross borders, I merge the GeoEPR with the Transnational Ethnic Kin (TEK) dataset, which matches each national sub-grouping with a larger international group, should one exist. The resulting matrix is 699x699.

The E weight matrix provides a basis for testing hypothesis (2), but is not sufficient for identification of the true effect of ethnicity. Because most ethnic groups are scattered across a close cluster of districts, parameter estimates might be capturing relationships of direct contiguity, rather than shared preferences and behaviors based upon a shared identity. The residual ethnicity matrix excludes four regions and has the dimensions 694x694.

All four types of neighborhoods provide the basis for evaluating hypothesis (3), that the mechanisms driving the spread of riots and protests and those causing violent conflict to spread are different. In particular, G and G_R should be more fundamental to violent conflict; E and E_R, to riots and protests.

Table 3 summarizes the four weight matrices. As the weight matrix becomes progressively smaller, the proportion of nonzero links rises, while both the average and maximum number of links per district fall. Under all four regimes, the most connected region is Kivu in eastern Congo: one of the most conflict-prone places in the entire world.

In the analysis, the row-stochastic form (each row normalized to sum to one) of each weight matrix is used. Normalization helps to ensure invertibility of the term $(I_N - \rho W)$- where ρ may be replaced by any of the three spatial weight parameters), and positive-definiteness of the variance-covariance matrix (LeSage, 2008).

4.3 Independent variables

Weight matrices are the most powerful and direct way of testing a spatial hypothesis. The final two hypotheses address the impact of political factors upon the diffusion process. Ideally, a political contiguity matrix could be constructed to test these two hypotheses. Such a matrix could be constructed in a number of ways, but none would be plausible nor suitable for evaluating the precise influence of political power upon conflict diffusion. Clear-cut political affiliations, such as pro-communist vs. pro-western governing philosophies, have lost their meaning since the end of the Cold War. Classifying regimes as autocracies or democracies is necessarily subjective; moreover, the type of government is frequently an endogenous outcome of conflict processes.[10] In all cases, ideological affiliations and types of political systems are fluid, and frequently fast-changing. The spatial weight matrix is instead assumed to be exogenous and constant over the entire period of study.

Hypothesis (4a) is that conflict shocks originating from districts with political power will be less vulnerable to conflict than will isolated or ostracized districts. Hypothesis (4b) posits that conflict shocks originating from powerful districts will provoke a more powerful response in terms of conflict spillovers. In order to evaluate these hypotheses, I introduce explanatories and interaction terms into regression equation that identify powerful and peripheral districts in a precise way.

The first two indicators measure a district's *access to central power*. The $capital_{it}$ variable equals 1 if the district hosts the national capital, and 0 otherwise. The $access_{it}$ variable equals 1 if either the current head of government or head of state was born in the district. Place of birth of the two most powerful state officials has economic and political significance because patronage networks are often targeted to benefit the specific ethnic group and region of origin of leaders (De Juan and Bank, 2015).[11] In Zimbabwe, for example,

[10] See Glaeser et al. (2004) for a similar argument about the endogeneity of institutions.
[11] Francois et al. (2015) finds that the extent of patronage is limited by the risk of outside rebellion, resulting in a fairly equal distribution of cabinet shares across different ethnic groups;. However, the country leader's ethnic group always receives a power

Robert Mugabe's own ethnic group- the Shona- enjoy special redistributive privileges, while the homeland of Zimbabwe's second largest ethnic group, Matabeleland, has been systematically ignored and neglected (Meredith, 2007).

The next indicators instead address *challenges to central power*. The $outside_{it}$ variable indicates whether the current head of state or government of a foreign nation was born in district i. This is a common occurrence due to the shift in international borders that followed decolonization in Africa, and is potentially indicative of a district with outside loyalties (though not necessarily one isolated from the government of its own state). For example, both presidents of Djibouti have been foreign-born, in Somalia and Ethiopia.

The $distance_{it}$ variable instead measures the distance (in km) of each district centroid from the national capital. The state's ability to exercise power is often deteriorating in this distance. As this measure grows, districts become more and more isolated from the apparatus of power, and less integrated into the nation as a whole. Therefore, it precisely indicates peripheral and marginalized districts, districts with a greater innate potential to challenge state power or even attempt to break away.

One drawback to the $distance_{it}$ and $capital_{it}$ indicators is that they present little variation over time. Over the entire period studied, two countries- Nigeria and Tanzania- moved their national capitals (in 1991 and 1996). Two more- Eritrea and South Sudan- became independent (in 1993 and 2011), which also registers as a transfer of the capital. [12]

By interacting the conflict variables with $capital_{it}$ (or $access_{it}$) on the one hand, and $outside_{it}$ and $distance_{it}$ on the other, we can precisely distinguish between those shocks originating from regions with access to, or peripheral from, central power. This distinction is important for addressing hypothesis (4b), that is that shocks coming from powerful districts and from peripheral ones will provoke a different response in terms of conflict behavior.

The set of capitals is taken from the *CIA World Factbook* (2016); the Latitude.to website, which calculates the GPS coordinates of any address in the world, provided their precise location. The two top politicians of each state are taken from the *World Factbook* and from the *World Statesmen* website, an online encyclopedia that catalogs the leaders of states and territories. Data on the districts of origin of politicians was obtained from various sources, including the websites of national parliaments, which often contain detailed biographies of MPs; international press agencies; and online newspapers.

4.4 Control variables

The control variables are nightlight data, population, population density, rainfall and deviations from average rainfall. These account for variation in how large and sophisticated locals economies are, and for the opportunity cost of fighting.

premium.

[12] Sovereignty over the capital of Western Sahara, Laayoune, is claimed by both Morocco and the Polisario Front, and the day-to-day activities of the independence movement are generally directed from refugee camps located elsewhere. Because of its lack of sovereignty, Western Sahara is counted as having no capital at all in the dataset. See the *CIA World Factbook* (2016).

Nightlight satellite data has been used as a proxy for local levels of development in numerous studies (Henderson et al., 2012; Bickenbach et al., 2013; Mellander, et al., 2013; Nordhaus and Chen, 2014). Shortland (2013) uses variation in nighlight output to gauge the impact of Somalia's civil war upon local economic development; De Juan and Bank (2015), as a proxy for selective redistribution. For this study, nightlight data is taken from NOAA's Version 4 DMSP-OLS Nighttime Lights Time Series. The Average Visible, Stable Lights, & Cloud Free Coverages dataset was used. The yearly average of nightlight data was calculated for each district.

Because nightlight satellite data may simply be proxying urbanization, I also include controls for district population and population density. These data are taken from NASA's Socioeconomic Data and Applications Center (SEDAC). This data is only available at five-year intervals. Data for 2005 and 2010 are taken from the Gridded Population of the World, v4 dataset, while data for 1995 and 2000 are taken from the Global Rural-Urban Mapping Project (GRUMP), v1. Observations from 1997 to 2000 are assigned the 1995 values; from 2001 to 2004, the 2000 values, and so forth. In all cases, the UN-adjusted population counts are used. The original population density variable is persons/km; due to large variability in the size of districts, I rescale by dividing by 1,000. I rescale the population variable by dividing by 100,000.

Many recent papers employ rainfall data from satellites and gauges to demonstrate how climatic volatility influences local livelihoods and hence, the risk that regions will be struck by violence. Rainfall measures are particularly suited to explaining conflict incidence in sub-Saharan Africa: The one region of the world where a majority of livelihoods still depend upon rain-fed agriculture. Environmental instability plays a much weaker role elsewhere (Miguel; Gleditsch, 2012).

While some work finds negative deviations in rainfall to be correlated with heightened conflict risk (Miguel; Harari and La Ferrara, 2013), other authors argue that droughts can induce cooperation between competing ethnic groups. Rainfall abundance may pose a greater risk (Gleditsch, 2012). There is some evidence that flooding is even more dangerous than drought, as it can disrupt livelihoods in more extreme ways and generate destructive patterns of scarcity (Ibid). Other work finds extreme rainfall patterns to be robustly associated with drops in growth, but not in conflict. Indeed, no link between conflict risk and scarcity of any kind of renewable resource– as characterized by drought, soil degradation, deforestation or population density?remains robust to alternate specifications and estimation strategies (Buhaig et al., 2008).

Because extreme deviations- positive and negative- from average monthly rainfall levels are more important to explaining conflict than are rainfall levels, I employ the former as a control in all specifications. Both measures are calculated on the basis of data from the CPC Merged Analysis of Precipitation (CMAP), available from the National Oceanic and Atmospheric Administration (NOAA) of the U.S. Department of Commerce.

4.5 Estimation and identification challenges

Individual effects, μ_i, are introduced into all three models in order to control for time-invariant heterogeneity among the observed regions. Under the fixed effects assumption, the dummy measures the intercept for each region. The random effects assumption instead specifies the intercept as a random variable, and does not allow the individual effect and the error term to be correlated. When the units of observation have irregular shapes and sizes, as in our case, the random effects assumption is inappropriate (Elhorst, 2003). Hausman tests to compare the two models confirm that the random effects assumption is inappropriate for this dataset; therefore, all models are estimated with fixed effects.

Fixed effects require the within transformation: Each variable is replaced by its deviation from the individual mean (ie, is time demeaned) and constant variables- including the individual effects μ_i- disappear from the model. The F.E. model is unbiased if there are no unobservables that vary across both time and space. I introduce year dummies to control for time-varying influences that are constant across individuals. A number of specification tests- including Wooldridge's test for unobserved effects in panel models, the Lagrange multiplier tests, and the F test- confirm that the correctly specified model requires both individual and time fixed effects.

The estimation process of the fixed effects spatial panel model is a two-step iterative procedure, alternating between GLS for the nuisance parameters (β and σ_ν^2) and maximization of the concentrated likelihood for the parameters of interest (the three spatial parameters, as well as the variance ratio $\phi = \sigma_\mu^2/\sigma_\varepsilon^2$). These two steps- GLS and ML- are repeated until a convergence criteria for parameter estimates is satisfied. See Millo and Piras (2012) and Croissant and Millo (2008) for further details on the estimation procedure used by plm and splm, the two packages in R dedicated to panel data and spatial panel data analysis.

Maximum likelihood estimation of the F.E. spatial panel model imposes a number of restrictions on the stochastic term, including normality, homoskedasticity, and the absence of serial correlation. Several of the assumptions prove to fail in this dataset. Applying the Breusch-Pagan test, the null hypothesis of homoskedasticity is rejected for several of the 18 cross-sections. Moreover, the Breusch-Godfrey test and the Wooldridge test for AR(1) errors, as well as the the Baltagi, Song, Jung and Koh LM test (specific to spatial panels) reveal the error terms to be serially correlated. The within transformation required by fixed effects leads to residuals with negative serial autocorrelation, but in this case the dependence applies to the original data as well.

One solution to serial correlation is to introduce time-lags of the dependent. However, when the fixed effects transformation is applied this leads to correlation between the regressors and the error term, leading to a bias (the "Nickell Bias") in the estimate of the SAR parameter of order 1/T (Nickell, 1981). For my panel dataset (18 years long) the Nickell Bias is of the order 0.06. Both the time and spatial lag coefficients are impacted. However, coefficient estimates are biased downwards. Therefore, specifications that include time-lagged dependents can be interpreted as a useful lower bound on the true effect.

When serial correlation and heteroskedasticity is present, standard errors tend to underestimate the true standard deviation, and therefore cannot be used a basis for statistical inference. Serial correlation also leads to inconsistent estimates in the SAR model. I allow for both serial and spatial autocorrelation in the disturbances of all spatial specifications. The simplified framework discussed above assumed that the non-spatially autocorrelated part of the error followed the distribution $\nu_{it} \sim IID\left(0, \sigma_\nu^2\right)$. In this paper, the

ν are instead taken be serially autocorrelated. The complete model for the disturbance, presented in Millo (2013), has an individual, time-invariant component μ (as before) and an idiosyncratic component that is both spatially and serially autocorrelated. The disturbance vector has three parts:

$u = (\iota_T \otimes \mu) + \varepsilon$

$\varepsilon = \rho \left(I_T \otimes W\right) \varepsilon + \nu$

$\nu_t = \psi \nu_{t-1} + e_t$

The new parameter ψ captures serial correlation.

4.5.1 The reflection problem

Employing the spatial lag of conflict as a dependent variables, as all of the models presented here do, presents a special challenge to identification. The reflection problem, introduced by Manski (1993), closely relates to the discussion of clustered risk factors versus true interdependence. When districts in a neighborhood engage in similar behavior, does this reflect the dependence of individual behavior upon the propensity of neighbors to adopt the same action (endogenous effects)? Or are similar districts responding to similar circumstances (correlated effects)? To paraphrase Manksi, when we observe both a district and its neighbor to engage in conflict simultaneously, is the influence of the neighbor responsible for the district's behavior, is the neighbor reflecting the district, or are both district and neighbor responding to a common shock?[13]

From a statistical point of view, the lagged dependents are by design endogenous regressors, correlated with the error term. However, when the functional form and spatial weight matrix are correctly specified, maximum likelihood provides consistent estimates of the true spillover effect.

5 Results

In this section, I present estimation results. OLS estimates of the panel data model are discussed in section 5.1, while the SAR model, estimated with ML, is discussed in 5.2. The Durbin model is presented in 5.3. In the following section, I introduce political interactions into the SAR model in order to better evaluate the impact of these variables. Finally, robustness checks are presented in section 5.5. Neighborhoods are restricted to only internal and external neighbors; a spatially autocorrelated (SEM) error term is introduced into the SAR model; and the models are estimated with different types of fixed effects (under construction).

5.1 OLS estimates of the non-spatial model

The OLS estimates provide a baseline for comparison with the later spatial models. The three dependents I model in Table 3 are total number of events, number of riots and number of violent events. In column (1),

[13]"Does the mirror cause the person's movements, does the image reflect the person's movements, or do the person and the image move together in response to a common external stimulus?' (Manski, 1996).

I regress events on four basic controls: the nightlight variable, population, population density and deviations from average monthly rainfall. In column(2), time-lagged events are introduced. As discussed above, this coefficient will be smaller in magnitude than the true value. In column (3), I regress riots on the set of basic controls, introducing time-lagged riots in column (4). The two violence regressions are presented in columns (4) and (5). All models are run on the 764x764 geographic contiguity dataset.

The population and the nightlights indicators are positively associated with all types of conflict: Larger, more economically developed regions experience more riots, more violence, and greater conflict overall. Excessive rainfall and flooding also leads to more conflict, but loses its significance when the lagged dependent is introduced. Finally, the coefficient on the time-lagged dependent is large and highly significant in all cases.

Distance is the most important of all the political variables. As the distance between the district and the national capital grows, the level of conflict it experiences goes down. Access to outside power leads to less violence (significant at the 10% level). Finally, access to central power leads to more conflict events of any type (significant at the 10% level). The capital dummy is not significant in any of the regressions.

Much as traditional specification tests evaluate whether the Gauss-Markov assumptions are met, OLS residuals can be also analyzed for sptatial autocorrelation. Moran's I varies between -1 (negative autocorrelation, so that like observations of the dependent repel each other) and +1 (positive autocorrelation, so that like observations cluster together). For each of the 18 cross-sections that make up the panal dataset, the residuals from specification (1) tested positive for spatial correlation, with an average Moran's I of 0.228 and an average p-value of 0.001.

Figure 2 maps out the residuals from the estimation of (2) on the entire panel. This map provides clear additional evidence of spatial clustering. Broad swaths of the continent display negative residuals, in blue. In green are regions for which the model underestimates violence, resulting in positive residuals: these include much of the Sahara, coastal Egypt and Sudan, the Horn of Africa and Madagascar.

5.2 SAR model estimates

Regression results for the SAR model are presented in Table 4. Lambda measures the average impact of spatial lags in the dependent upon a district's own outcome. It is positive and significant at the 1% level across all specifications.

Panel A presents estimates based upon the contiguity weights matrix, G. Rioting displays the greatest spatial autocorrelation of any type of conflict, although the magnitude of the coefficient falls when time-lagged riots are introduced (a conservative underestimate, given the Nickell Bias). This holds true for all types of conflict. Spatial autocorrelation is much lower in the residual contiguity model (panel B), suggesting that the effect of geography was confounded by shared ethnic ties.

Ethnicity is an equally important source of conflict spillovers, as shown in panel C. Regions that share ethnic groups are particularly vulnerable to the spillover of rioting and protests. Even when ethnicity is purged of the effect of direct geographical ties (panel D), riots display a high degree of spatial autocorrelation.

These results in Table 4 confirm the first two hypotheses. Both geography and ethnicity are important channels for the diffusion of conflict. Observed patterns of violence across districts are not independent, but the outbreak in one district actually discourages its spread elsewhere. When conflict intensifies in one district, there is a positive neighborhood effect.

Hypothesis (3) posited that spillovers of violent conflict would be more significant for those neighborhoods defined by geographical contiguity. This is not what we find. We find, in fact, that violence spreads more easily through ethnic networks. Ethnicity is more important than geography for the diffusion of all types of conflict. Yet its impact is particularly strong when we consider the sort of conflict that spreads through soft, indirect channels: Riots and protests.

The rioting SAR coefficient for the neighborhood of districts defined by ethnicity, equal to 0.62, has two possible interpretations. First, when the average number or riots in ethnically-linked districts increases by one, the number of riots in the average district will increase by 0.62. Second, when the number of riots in district i increases by one, the average district j will undergo a riot shock of 0.62/N, where N is the number of j's ethnic neighbors. In the same panel, we find that when the weighted average of violent events increases marginally in an ethnic neighborhood, the average district undergoes a violence shock equal to 0.41.

The four political indicators are introduced into the contiguity and residual contiguity SAR regressions in Table 5a, and into the ethnicity and residual ethnicity regressions in Table 5b. The addition of the political variables does not modify the magnitude of the SAR coefficients; nor does their statistic significance fall. In the contiguity models, distance from the capital is associated with a slight reduction in violent events and in overall conflicts, much as we found in the OLS model with political variables. This effect is significant only in the regressions with time-lagged dependents, however. Distance exercises a similarly negative impact upon conflict intensity in the two ethnicity models. In the ethnicity model, for example, we find that an increase in distance from the capital of 100 km is associated with two less violent events for the average district, and two fewer incidents of conflict of any kind.

There is weak evidence that access to outside power increases rioting (in the ethnicity model), while decreasing violence (in the residual ethnicity model). But this result is not robust.

5.3 Spatial Durbin model estimates

Further evidence for the impact of the political variables is provided by the spatial Durbin models (Table 6a-6d). Each of the four tables presents Durbin estimates for a different set of neighborhoods: Contiguity and residual contiguity in 6a and 6b, ethnicity and residual ethnicity in 6c and 6d. In each table, the dependents are number of conflict events (column 1), number of riots (column 2), and number of violent events (column 3). The time-lagged dependent is not included in the regressions.

In all regressions, the SAR parameter maintains the same magnitude as in previous regressions, and remains highly significant. However, the combination of spatial autocorrelation in the dependent and independent variables performs quite poorly for one dependent, in particular: violence. With an adjusted R-squared equal to 0 in all of the violence regressions, it is clear that the Spatial Durbin is not at all able to explain the variation across time of this outcome. It performs slightly better when number of events and rioting are the modelled outcomes. Moreover, the lagged independents are never significant for the two contiguity models. This result suggests that neighborhood effects are far more powerful and important for those neighborhoods defined by ethnic ties, rather than geographical nearness.

The Durbin coefficient, given by γ, represents the indirect effect of marginal change in a indicator; that is, the average impact upon a district's neighbors of policies the district chooses to enact, or of shocks it undergoes. In the ethnicity and residual ethnicity models, three of the Durbin coefficients are statistically significant at the 5% level. Of the spatially lagged controls, only population and population density are significant, while rainfall deviations are significant at the 10% level.

When the weighted average of population in an ethnic neighborhood increases by 100,000, the average district experiences about one more incident of violent conflict. The increase is 1.33 for the residual ethnicity neighborhood. Thus there is a positive association between population and violence. However, population density has the opposite effect. An increase of 1000 residents/km in the average density of an ethnic neighborhood leads to 13.20 fewer conflict events for the typical region. Alternatively, when population density for one district increases by 1000 residents/km, all districts in the neighborhood experience a decrease of 13.20/N conflict events, where N is average number of links. Although these shocks to population and population density may seem unrealistically large, half a million Hutus fled from Rwanda to Eastern Congo in the wake of the 1994 genocide. The UNHCR estimates that 1.1 million Syrians and 450,000 Palestinians were registered as refugees in Lebanon in 2014, a country of only 4.1 million inhabitants.

Of the spatially lagged political indicators, access to central power is positively and significantly related to number of riots. When a native inhabitant of an ethnically neighboring region becomes president or prime minister, the network as a whole experiences 3.43 more riots per year, on average.

5.4 At the intersection of conflict and political power

The SAR and spatial Durbin models provide mixed evidence for hypotheses (4a) and (4b). There is little evidence that more peripheral regions are more vulnerable conflict; in fact, the opposite appears to hold true, as regions more distant from the capital experience lower levels of conflict. Conflict in capital regions does not appear to generate strong spillovers; nor is access to power in an outside district of particular importance. However, spillovers from districts with access to central power do appear to be of some importance to the frequency of rioting.

In this section, I introduce interactions between the political variables and time-lagged conflict outcomes. In Tables 7a and 7b, each outcome is interacted with the capital dummy, access to central power, access to outside power, and distance from the capital. The dependent variable in each case is number of conflict events. The capital coefficient is positive and highly significant across all three specifications in which it is included. The interaction with rioting is negative and significant at the 10% level. Access to central power and its interactions are not significant.

Although the outside dummy is not significant, each of its interactions is negative and significant: Obtaining access to outside power- in terms of a native inhabitant of the region acceding to rule in a foreign country- depresses number of conflict events in the following year, when this access to outside power is paired with an increase in conflict in the current year.

Finally, distance from the capital is consistently negative and highly significant, as previous regressions have shown. However, an increase in distance from the capital does not depress future conflict when it is combined with a higher frequency of conflict in the current year.

Rioting is the outcome in Tables 7c and 7d. Only the interaction between rioting and distance from the capital is significant (and negative).

Finally, violence is the outcome in Tables 7e and 7f. There is a strong positive association between the capital dummy and a violence, though none of the interaction terms are significant. All of the outside interactions are negative and significant, and once again distance is negative and significant. Thus the violence estimates mirror those obtained for overall conflict.

The political interactions suggest that the capital district is uniquely vulnerable. While conflict shocks originating from the capital do not seem to significantly influence conflict, it is regions most exposed to central power- not the most isolated and peripheral- that are subject to the greatest amount of conflict. The farther a district is from the central capital, the less vulnerable it is to conflict. Finally, access to power in a foreign country, combined with past experiences of conflict, appear to depress future levels of conflict.

6 Robustness Checks

In this section, I perform several robustness checks. First, I allow for spatial autocorrelation of the error component, and estimate a combined SAR-SEM model. This model allows us to account for latent variables that may be spatially dependent and correlated with conflict. Second, I restrict each type of neighborhood to the set of internal and external neighbors, in order to isolate how conflict spreads within individual countries and across international borders. Finally, I reestimate the SAR model with different types of fixed effects in order to show that the results presented in the main part of the paper are robust (under construction).

The SAC model results are presented for the contiguity and residual contiguity networks in Table 8a, and for the ethnicity and residual ethnicity networks in Table 8b. In each case, the SAR parameters for number of events and for violence grow drastically in magnitude, while the SAR parameter for riots grows much smaller. Spillovers effects remain strongest for those neighborhoods defined by ethnicity. Finally, the SAR parameter becomes very small and loses its significance in those specifications that include the time-lagged dependent; yet given the downward Nickell Bias, these results should be taken with a grain of salt.

The SAR parameter appears to become smaller in the case of rioting because a large amount of positive spatial autocorrelation is picked up the SEM parameter. Conversely, the event and violence SAR parameters grow stronger than before because they are no longer offset by a large, negative SEM parameter. Overall conflict and violent incidents, on the one hand, and rioting on the other appear to be driven by quite different spatial processes. While overall events and violence demonstrate strong diffusion effects, much of the spatial correlation in rioting may in fact be driven by that of latent variables. Finally, once the SEM component is introduced, rioting appears to demonstrate larger spillover effects in the residual ethnicity network than in the original ethnicity one- an unexpected and somewhat bizarre result.

Tables 9a-9d present the next set of robustness check results, relative to the internally and externally defined neighbors. Internal and external contiguity results are presented in Table 9a, the residual results in 9b, the ethnicity results in 9c and the residual ethnicity results in 9d.

Internal contiguity and the residual counterpart continue to display significant spatial autocorrelation, while external contiguity and its residual counterpart do not. The contiguity results we saw in previous tables were almost entirely driven by internal links. However, the external results are of interest because a number of political variables become significant. The capital district and access to outside power are both positively associated with conflict, while distance from the capital is negatively associated with it.

Neighborhoods defined by internal ethnic and residual internal ethnic links continue to demonstrate strong spatial correlations. The two external ethnicity SAR parameters are extremely small and significant only for riots and violence; in the case of riots, the parameter turns negative. As in the contiguity case, the positive spatial correlation of ethnic neighborhoods was driven by internal neighbors. Finally, distance from the capital continues to lower the intensity of conflict experienced by a conflict.

7 Conclusions

The evidence presented in this paper supports two conclusions. First, capital districts are exposed to greater amounts of conflict. The farther a district is from the capital, the safer it becomes. However, it is not clear that conflict shocks originating from the capital are more dangerous to neighbors.

Second, both geographic and ethnic ties are important conduits for diffusion, but ethnicity is a more powerful force than geography.

Ethnicity appears to convey not only a shared sense of identity, but moreover new information about the desirability of engaging in conflict, and particularly in nonviolent conflict directed against the political system. A possible realignment of incentives is occurring. While refugee flows are by far the most visible, controversial and contentious consequence of modern warfare, these flows are frequently directed at the nearest international border. The degree to which refugees, combatants and arms eventually do succeed in finding their way to noncontiguous regions inhabited by ethnic kin is an open question. It is clear that neighborhoods defined by ethnicity, as opposed to geography, are engaging in an altogether different sort of calculus. Conflict spillovers are transforming these neighborhoods in profound and far-reaching ways. Spillovers from ethnic neighbors are- for better or for worse- destabilizing existing political systems,and remaking societies;

remaking, even, how societies see themselves.

FIGURE 1.

The contiguity (green), ethnicity (grey) and residual contiguity (purple) weight matrices

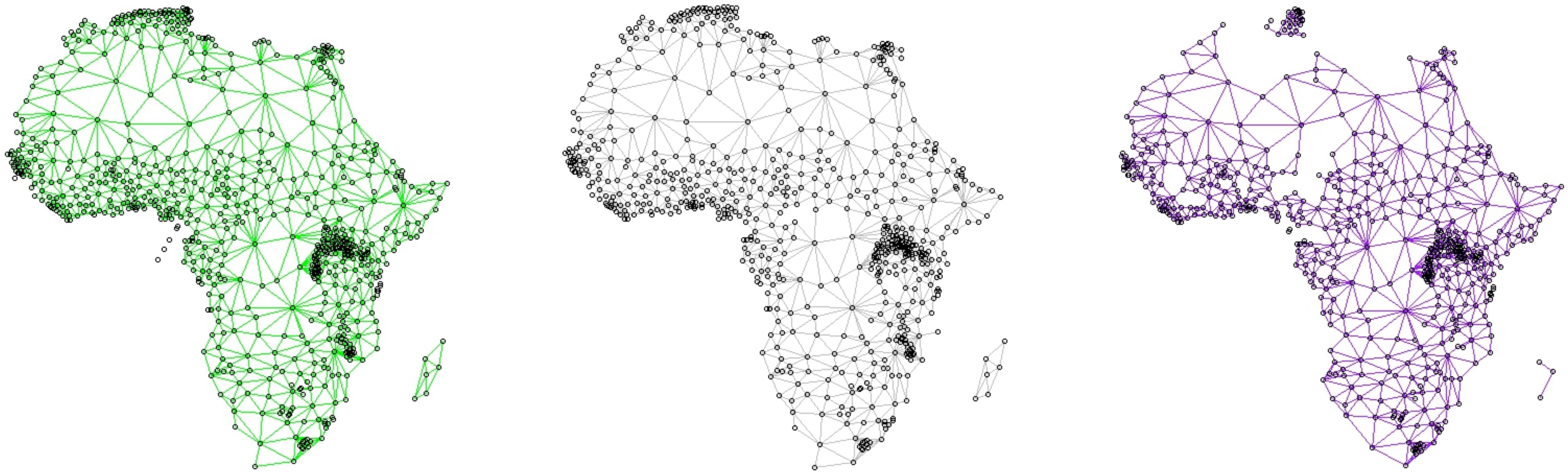

FIGURE 2.

Spatial residuals from the OLS model

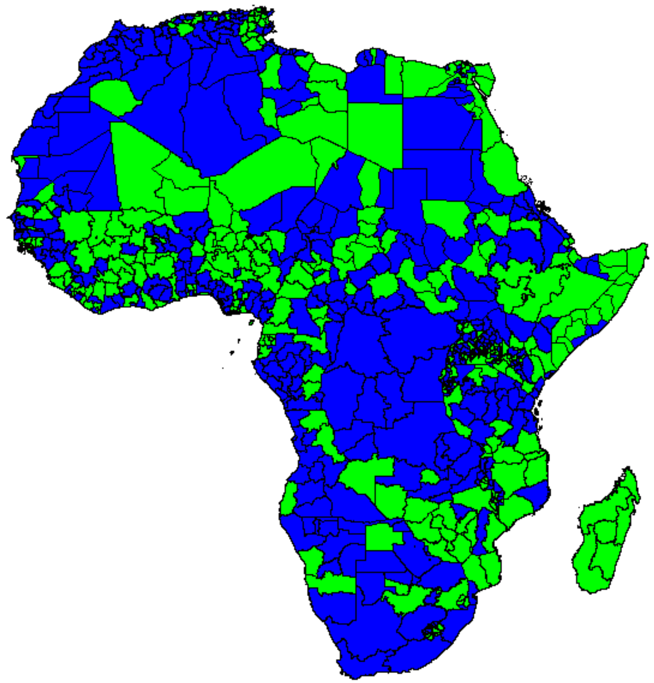

TABLE 1.

Descriptive statistics

Statistic	Mean	St. Dev.	Min	Max
Number of Events	8.27	34.35	0	879
Riots	2.09	12.07	0	652
Violent Events	5.50	26.41	0	785
Fatalities	42.10	900.85	0	66,565
Population	1,083,917	1,679,260	9.12	23,628,270
Population Density	247.59	882.37	0.08	17,268.71
Rainfall	3.57	3.67	0	25.17
Deviations in Rainfall	-1.46	1.97	-6.47	5.22
Nightlights	5.39	7.18	0	62.35
Capital	0.07	0.25	0	1
Access to Central Power	0.13	0.33	0	1
Access to Outside Power	0.01	0.08	0	1
Distance from Capital	321.77	266.25	1.03	1,436.06

TABLE 2.

The four spatial networks

	E	E_R	G	G_R
N	699	694	764	592
% nonzero links	2.49	2.07	0.7	0.53
ave. number of links	17.4	14.34	5.37	3.14
max. number of links	95	64	18	15

TABLE 3.

OLS model: Contiguity

	(1) N. Events	(2) N. Events	(3) Riots	(4) Riots	(5) Violence	(6) Violence
Population	1.61**	0.67**	0.96**	0.41*	0.56+	0.22+
	(0.56)	(0.24)	(0.36)	(0.19)	(0.29)	(0.12)
Pop. density	2.33	-0.61	-0.71	-0.72	2.62	0.07
	(7.39)	(2.44)	(1.68)	(0.67)	(6.27)	(2.09)
Nightlights	1.29***	0.72***	0.74***	0.36***	0.45*	0.34**
	(0.28)	(0.15)	(0.12)	(0.07)	(0.19)	(0.11)
Rainfall deviations	0.68***	0.00	0.21***	0.02	0.54***	0.07
	(0.20)	(0.14)	(0.05)	(0.04)	(0.16)	(0.12)
Capital	-20.03	-6.97	-1.21	-0.89	-22.86	-7.79
	(20.91)	(7.78)	(2.69)	(1.24)	(18.09)	(6.66)
Access	3.20	1.93+	1.10	0.70	1.58	0.91
	(2.24)	(1.16)	(1.30)	(0.63)	(1.30)	(0.75)
Outside	-3.39	-1.59	-1.32	-0.16	-2.00	-1.62+
	(2.83)	(1.04)	(1.30)	(0.68)	(1.63)	(0.92)
Distance	-0.06*	-0.03***	-0.00+	-0.00*	-0.05*	-0.03***
	(0.03)	(0.01)	(0.00)	(0.00)	(0.02)	(0.01)
Lagged dependent		0.71***		0.68***		0.70***
		(0.03)		(0.09)		(0.05)
Constant	5.29	3.20	-10.86*	-4.56*	15.53+	7.65*
	(11.66)	(4.39)	(4.34)	(2.24)	(9.02)	(3.49)
R-squared	0.03	0.48	0.06	0.43	0.01	0.46
N	764	764	764	764	764	764

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

TABLE 4.

SAR Models

	(1) N. Events	(2) N. Events	(3) Riots	(4) Riots	(5) Violence	(6) Violence
Contiguity						
Lambda	0.40***	0.22***	0.48***	0.36***	0.43***	0.25***
	(0.05)	(0.03)	(0.08)	(0.06)	(0.07)	(0.04)
Lagged dependent		0.65***		0.62***		0.64***
		(0.04)		(0.08)		(0.06)
R-squared	0.04	0.57	0.07	0.39	0.02	0.63
N	764	764	764	764	764	764
Resid. Contiguity						
Lambda	0.25***	0.13***	0.26***	0.20***	0.25***	0.13***
	(0.05)	(0.03)	(0.05)	(0.04)	(0.05)	(0.03)
Lagged dependent		0.69***		0.80***		0.67***
		(0.04)		(0.08)		(0.06)
R-squared	0.04	0.61	0.08	0.58	0.04	0.64
N	592	592	592	592	592	592
Ethnicity						
Lambda	0.41***	0.25***	0.62***	0.46***	0.41***	0.26***
	(0.05)	(0.04)	(0.10)	(0.08)	(0.05)	(0.04)
Lagged dependent		0.64***		0.59***		0.61***
		(0.03)		(0.07)		(0.04)
R-squared	0.06	0.55	0.08	0.41	0.01	0.58
N	699	699	699	699	699	699
Resid. Ethnicity						
Lambda	0.31***	0.20***	0.52***	0.38***	0.27***	0.17***
	(0.04)	(0.03)	(0.10)	(0.07)	(0.04)	(0.02)
Lagged dependent		0.64***		0.62***		0.60***
		(0.03)		(0.08)		(0.04)
R-squared	0.06	0.55	0.08	0.42	0.02	0.57
N	694	694	694	694	694	694

$^{+}\ p < 0.10$, $^{*}\ p < 0.05$, $^{**}\ p < 0.01$, $^{***}\ p < 0.001$

TABLE 5a.

SAR Models with political variables: contiguity and residual contiguity

	(1) N. Events	(2) N. Events	(3) Riots	(4) Riots	(5) Violence	(6) Violence
Contiguity						
Lambda	0.40***	0.22***	0.48***	0.36***	0.43***	0.25***
	(0.05)	(0.03)	(0.08)	(0.06)	(0.07)	(0.04)
Capital	-4.44	0.01	1.50	0.67	-10.88	-2.88
	(16.86)	(6.65)	(3.03)	(1.45)	(14.15)	(5.55)
Access	2.85	1.82	1.10	0.73	1.18	0.74
	(2.01)	(1.13)	(1.19)	(0.60)	(1.11)	(0.71)
Outside	-2.92	-1.41	-0.90	0.03	-2.11	-1.69
	(2.81)	(1.18)	(1.38)	(0.65)	(1.83)	(1.21)
Distance	-0.03	-0.02*	0.00	0.00	-0.03	-0.02*
	(0.02)	(0.01)	(0.00)	(0.00)	(0.02)	(0.01)
Lagged dependent		0.65***		0.62***		0.64***
		(0.04)		(0.08)		(0.06)
R-squared	0.03	0.53	0.07	0.39	0.00	0.55
N	764	764	764	764	764	764
Resid. Contiguity						
Lambda	0.25***	0.13***	0.26***	0.20***	0.25***	0.13***
	(0.05)	(0.03)	(0.05)	(0.04)	(0.05)	(0.03)
Access	1.72	1.27	0.60	0.45	0.38	0.27
	(2.34)	(1.24)	(1.51)	(0.57)	(1.12)	(0.68)
Outside	-1.04	-0.42	-1.23	-0.15	-0.51	-0.60
	(2.86)	(1.62)	(1.79)	(0.76)	(1.94)	(1.27)
Distance	-0.04	-0.02*	0.00	0.00	-0.04+	-0.02**
	(0.02)	(0.01)	(0.00)	(0.00)	(0.02)	(0.01)
Lagged Dependent		0.68***		0.80***		0.67***
		(0.04)		(0.08)		(0.06)
R-squared	0.02	0.57	0.08	0.58	0.00	0.57
N	592	592	592	592	592	592

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

TABLE 5b.

SAR Models with political variables, ethnicity and residual ethnicity

	(1) N. Events	(2) N. Events	(3) Riots	(4) Riots	(5) Violence	(6) Violence
Ethnicity						
Lambda	0.41^{***}	0.25^{***}	0.62^{***}	0.46^{***}	0.40^{***}	0.26^{***}
	(0.05)	(0.04)	(0.10)	(0.08)	(0.05)	(0.04)
Capital	-4.62	0.53	1.66	0.86	-11.24	-2.79
	(18.21)	(7.49)	(3.06)	(1.48)	(15.86)	(6.71)
Access	3.05	1.62	0.59	0.33	1.73	0.86
	(2.12)	(1.24)	(1.17)	(0.58)	(1.17)	(0.81)
Outside	-2.98	-1.46	0.07	0.73^{+}	-2.79	-2.19
	(2.61)	(1.48)	(0.65)	(0.44)	(1.98)	(1.49)
Distance	-0.03	-0.02^{*}	-0.00	-0.00	-0.03	-0.02^{+}
	(0.02)	(0.01)	(0.00)	(0.00)	(0.02)	(0.01)
Lagged dependents		0.64^{***}		0.59^{***}		0.61^{***}
		(0.03)		(0.07)		(0.04)
R-squared	0.04	0.50	0.08	0.41	0.00	0.49
N	699	699	699	699	699	699
Resid. Ethnicity						
Lambda	0.31^{***}	0.20^{***}	0.52^{***}	0.38^{***}	0.26^{***}	0.16^{***}
	(0.04)	(0.03)	(0.10)	(0.07)	(0.04)	(0.02)
Capital	-14.18	-5.22	0.89	0.27	-19.48	-8.28
	(19.98)	(8.56)	(3.03)	(1.39)	(17.62)	(8.00)
Access	3.01	1.59	0.61	0.34	1.76	0.89
	(2.22)	(1.27)	(1.25)	(0.61)	(1.23)	(0.83)
Outside	-4.25	-2.12	-0.55	0.52	-3.10^{+}	-2.41
	(2.60)	(1.56)	(0.77)	(0.51)	(1.83)	(1.50)
Distance	-0.04	-0.02^{*}	0.00	-0.00	-0.04^{+}	-0.02^{*}
	(0.03)	(0.01)	(0.00)	(0.00)	(0.02)	(0.01)
Lagged Dependent		0.64^{***}		0.62^{***}		0.60^{***}
		(0.03)		(0.08)		(0.04)
R-squared	0.04	0.49	0.08	0.42	0.00	0.44
N	694	694	694	694	694	694

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

TABLE 6a.

Durbin model: contiguity

	(1) N. Events	(2) Riots	(3) Violence
Contiguity- X			
Lambda	0.40***	0.48***	0.43***
	(0.05)	(0.08)	(0.07)
Population	1.21*	0.89*	-0.05
	(0.61)	(0.39)	(0.26)
Pop. density	3.22	0.15	2.51
	(5.89)	(1.57)	(4.43)
Nightlights	0.55	0.44**	-0.09
	(0.37)	(0.15)	(0.29)
Rainfall deviations	-0.83	-0.05	-0.70
	(1.32)	(0.20)	(1.18)
Capital	-5.17	-0.24	-9.37
	(15.96)	(3.39)	(12.47)
Access	2.72	1.07	0.76
	(2.00)	(1.19)	(1.07)
Outside	-3.10	-0.72	-2.18
	(2.85)	(1.29)	(1.81)
Distance	-0.03	-0.00	-0.02
	(0.02)	(0.00)	(0.02)
Wx			
Population	-0.09	-0.19	0.11
	(0.37)	(0.22)	(0.23)
Density	11.48	-0.58	11.33
	(19.11)	(4.88)	(17.16)
Nightlights	-0.34	0.13	-0.45
	(0.62)	(0.34)	(0.42)
Rainfall deviations	1.08	0.06	0.96
	(1.48)	(0.21)	(1.31)
Capital	-18.87	0.44	-9.90
	(41.99)	(8.02)	(36.03)
Access	-1.60	-1.83	0.65
	(3.48)	(1.94)	(2.22)
Outside	9.78	4.81	2.38
	(6.99)	(3.44)	(4.55)
Distance	-0.01	0.01	-0.02
	(0.03)	(0.01)	(0.03)
R-squared	0.02	0.07	0.00
N	764	764	764

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

TABLE 6b.

Durbin model: residual contiguity

	(1) N. Events	(2) Riots	(3) Violence
Resid. Contiguity- X			
Lambda	0.25***	0.26***	0.26***
	(0.05)	(0.05)	(0.05)
Population	1.19*	0.58**	0.24
	(0.54)	(0.19)	(0.40)
Pop. density	6.03	-4.78+	13.13
	(22.06)	(2.46)	(19.35)
Nightlights	0.22	0.66*	-0.65
	(0.68)	(0.33)	(0.60)
Rainfall deviations	-0.27	0.10	-0.35
	(0.98)	(0.12)	(0.90)
Capital	-8.15	2.48	-16.07
	(21.60)	(2.73)	(17.89)
Access	1.65	0.60	0.07
	(2.31)	(1.50)	(1.10)
Outside	-0.96	-1.07	-0.19
	(2.75)	(1.68)	(1.92)
Distance	-0.04	0.00	-0.04
	(0.03)	(0.00)	(0.03)
Wx			
Population	-0.61	-0.16	-0.29
	(0.43)	(0.17)	(0.33)
Pop. density	19.33	-3.76	22.70
	(37.66)	(2.99)	(36.09)
Nightlights	0.07	0.23	-0.26
	(1.00)	(0.52)	(0.68)
Rainfall deviations	0.66	-0.14	0.84
	(1.11)	(0.14)	(1.00)
Capital	28.20	-1.04	33.99
	(41.71)	(5.25)	(34.57)
Access	0.94	-0.06	1.12
	(2.05)	(0.77)	(1.57)
Outside	2.11	0.78	0.66
	(2.38)	(1.54)	(1.35)
Distance	0.01	0.00	0.01
	(0.03)	(0.00)	(0.02)
R-squared	0.02	0.09	0.00
N	592	592	592

$^{+}p < 0.10$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

TABLE 6c.

Durbin model: ethnicity

	(1) N. Events	(2) Riots	(3) Violence
Ethnicity- X			
Lambda	0.40***	0.61***	0.39***
	(0.05)	(0.10)	(0.05)
Population	0.91	0.80+	-0.22
	(0.68)	(0.43)	(0.32)
Pop. density	10.91	2.36	6.88
	(9.39)	(2.00)	(7.90)
Nightlights	0.16	0.24+	-0.19
	(0.47)	(0.13)	(0.38)
Rainfall deviations	0.52	-0.08	0.54
	(0.47)	(0.12)	(0.38)
Capital	-5.54	1.20	-12.25
	(17.69)	(3.06)	(15.19)
Access	3.03	0.66	1.49
	(2.10)	(1.16)	(1.15)
Outside	-2.48	0.09	-2.63
	(2.58)	(0.56)	(1.94)
Distance	-0.03	-0.00	-0.02
	(0.02)	(0.00)	(0.02)
Wx			
Population	1.06	-0.12	1.05+
	(0.67)	(0.21)	(0.55)
Pop. density	-13.20*	-5.76*	-4.30
	(6.19)	(2.65)	(4.45)
Nightlights	0.52	0.31	-0.25
	(0.80)	(0.58)	(0.28)
Rainfall deviations	-0.75	0.08	-0.68
	(0.61)	(0.15)	(0.49)
Capital	-69.30	-3.77	-61.54
	(75.75)	(10.46)	(69.33)
Access	3.81	3.43*	1.57
	(4.44)	(1.56)	(2.41)
Outside	30.76	1.24	29.39
	(22.24)	(5.62)	(19.00)
Distance	-0.02	0.01	-0.03
	(0.03)	(0.00)	(0.03)
R-squared	0.03	0.09	0.00
N	699	699	699

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

TABLE 6d.

Durbin model: residual ethnicity

	(1) N. Events	(2) Riots	(3) Violence
Resid. Ethnicity- X			
Lambda	0.29***	0.52***	0.24***
	(0.04)	(0.10)	(0.04)
Population	0.99	0.82+	0.05
	(0.69)	(0.44)	(0.34)
Pop. density	10.88	2.37	7.43
	(9.45)	(1.97)	(8.11)
Nightlights	0.17	0.22	-0.12
	(0.46)	(0.14)	(0.38)
Rainfall deviations	0.54	-0.04	0.51+
	(0.37)	(0.11)	(0.30)
Capital	-13.34	0.61	-18.48
	(19.78)	(2.85)	(17.44)
Access	2.85	0.62	1.60
	(2.21)	(1.25)	(1.23)
Outside	-3.36	-0.27	-2.81
	(2.43)	(0.55)	(1.80)
Distance	-0.03	0.00	-0.03
	(0.03)	(0.00)	(0.02)
Wx			
Population	1.29+	-0.17	1.33*
	(0.73)	(0.23)	(0.62)
Pop. density	-13.19**	-5.05**	-6.04+
	(4.88)	(1.87)	(3.30)
Nightlights	0.77	0.57	-0.16
	(0.74)	(0.53)	(0.28)
Rainfall deviations	-0.92+	0.02	-0.72+
	(0.50)	(0.14)	(0.37)
Capital	49.17	-0.99	49.89
	(86.46)	(7.39)	(79.69)
Access	3.16	2.54	1.80
	(4.00)	(1.85)	(2.34)
Outside	12.71	-5.28	19.39
	(14.82)	(4.12)	(12.85)
Distance	-0.02	-0.00	-0.03
	(0.03)	(0.01)	(0.02)
R-squared	0.04	0.09	0.00
N	694	694	694

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

TABLE 7a.

Political interactions: number of events

	(1)	(2)	(3)	(4)	(5)	(6)
			N. Events			
Lambda	0.25^{***}	0.35^{***}	0.30^{***}	0.25^{***}	0.36^{***}	0.31^{***}
	(0.04)	(0.04)	(0.05)	(0.04)	(0.04)	(0.05)
Lagged n. events	0.65^{***}			0.61^{***}		
	(0.04)			(0.06)		
Lagged riots		1.13^{***}			0.87^{***}	
		(0.18)			(0.08)	
Lagged violence			0.78^{***}			0.70^{***}
			(0.08)			(0.08)
Capital	17.54^{***}	23.78^{***}	22.12^{***}			
	(2.46)	(2.30)	(5.42)			
Capital x n. events	-0.04					
	(0.05)					
Capital x riots		-0.33^{+}				
		(0.20)				
Capital x violence			-0.11			
			(0.12)			
Access				1.25	2.68	1.21
				(1.05)	(1.64)	(1.61)
Access x n. events				0.04		
				(0.10)		
Access x riots					0.01	
					(0.16)	
Access x violence						0.12
						(0.21)
R-squared	0.53	0.15	0.35	0.55	0.17	0.36
N	699	699	699	699	699	699

$^{+}p < 0.10$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

TABLE 7b.

Political interactions: number of events, continued

	(1)	(2)	(3) N. Events	(4)	(5)	(6)
Lambda	0.25^{***}	0.36^{***}	0.30^{***}	0.25^{***}	0.34^{***}	0.30^{***}
	(0.04)	(0.04)	(0.05)	(0.04)	(0.04)	(0.05)
Lagged n. events	0.64^{***}			0.49^{***}		
	(0.03)			(0.06)		
Lagged riots		0.88^{***}			-0.11	
		(0.11)			(0.37)	
Lagged violence			0.76^{***}			0.57^{***}
			(0.07)			(0.13)
Outside	0.94	-0.54	0.10			
	(0.97)	(1.46)	(1.76)			
Outside x n. events	-0.35^{***}					
	(0.04)					
Outside x riots		-0.62^{*}				
		(0.28)				
Outside x violence			-0.56^{***}			
			(0.09)			
Distance				-9.42^{***}	-15.82^{*}	-8.88^{***}
				(2.50)	(6.66)	(2.41)
Distance x n. events				0.03^{*}		
				(0.01)		
Distance x riots					0.24^{*}	
					(0.10)	
Distance x violence						0.03^{+}
						(0.02)
R-squared	0.55	0.16	0.36	0.44	0.08	0.30
N	699	699	699	699	699	699

$^{+}p < 0.10$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

TABLE 7c.

Political interactions: riots

	(1)	(2)	(3)	(4)	(5)	(6)
			Riots			
Lambda	0.58***	0.46***	0.61***	0.59***	0.45***	0.62***
	(0.09)	(0.08)	(0.10)	(0.09)	(0.08)	(0.10)
Lagged n. events	0.05**			0.09**		
	(0.02)			(0.03)		
Lagged riots		0.58***			0.60***	
		(0.10)			(0.09)	
Lagged violence			0.03**			0.03**
			(0.01)			(0.01)
Capital	-7.33	1.65^{+}	0.18			
	(5.15)	(0.91)	(3.82)			
Capital x events	0.21*					
	(0.11)					
Capital x riots		0.01				
		(0.13)				
Capital x violence			0.06			
			(0.08)			
Access				0.05	0.37	0.31
				(0.76)	(0.52)	(1.08)
Access x n. events				0.03		
				(0.06)		
Access x riots					-0.01	
					(0.12)	
Access x violence						0.03
						(0.03)
R-squared	0.15	0.42	0.09	0.13	0.41	0.09
N	699	699	699	699	699	699

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

TABLE 7d.

Political interactions: riots, continued

	(1)	(2)	(3)	(4)	(5)	(6)
			Riots			
Lambda	0.59***	0.46***	0.62***	0.59***	0.43***	0.62***
	(0.10)	(0.08)	(0.10)	(0.10)	(0.08)	(0.10)
Lagged n. events	0.11**			0.25		
	(0.04)			(0.16)		
Lagged riots		0.59***			0.30+	
		(0.07)			(0.16)	
Lagged violence			0.04**			0.04
			(0.01)			(0.06)
Outside	0.43	0.74	0.33			
	(0.55)	(0.45)	(0.71)			
Outside x n. events	-0.04					
	(0.05)					
Outside x riots		-0.00				
		(0.12)				
Outside x violence			-0.05			
			(0.04)			
Distance				1.30	-0.76*	-0.17
				(1.38)	(0.35)	(0.90)
Distance x n. events				-0.03		
				(0.03)		
Distance x riots					0.07*	
					(0.03)	
Distance x violence						0.00
						(0.01)
R-squared	0.13	0.41	0.09	0.13	0.41	0.09
N	699	699	699	699	699	699

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

TABLE 7e.

Political interactions: violence

	(1)	(2)	(3)	(4)	(5)	(6)
			Violence			
Lambda	0.28***	0.40***	0.26***	0.27***	0.41***	0.26***
	(0.04)	(0.05)	(0.04)	(0.04)	(0.05)	(0.04)
Lagged n. events	0.44***			0.45***		
	(0.05)			(0.07)		
Lagged riots		0.45**			0.20***	
		(0.16)			(0.06)	
Lagged violence			0.62***			0.62***
			(0.05)			(0.08)
Capital	13.90**	12.99***	12.51***			
	(4.71)	(1.85)	(1.68)			
Capital x n. events	-0.13					
	(0.12)					
Capital x riots		-0.30+				
		(0.16)				
Capital x violence			-0.06			
			(0.05)			
Access				1.78+	1.57	1.10
				(0.92)	(1.11)	(0.76)
Access x n. events				-0.08		
				(0.08)		
Access x riots					0.04	
					(0.11)	
Access x violence						-0.03
						(0.13)
R-squared	0.39	0.02	0.54	0.42	0.01	0.58
N	699	699	699	699	699	699

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

TABLE 7f.

Political interactions: violence, continued

	(1)	(2)	(3) Violence	(4)	(5)	(6)
Lambda	0.29***	0.40***	0.27***	0.28***	0.39***	0.26***
	(0.04)	(0.05)	(0.04)	(0.04)	(0.05)	(0.04)
Lagged n. events	0.41***			0.15		
	(0.05)			(0.22)		
Lagged riots		0.23***			-0.40	
		(0.06)			(0.33)	
Lagged violence			0.61***			0.46***
			(0.04)			(0.06)
Outside	0.16	-1.48	-0.10			
	(1.13)	(1.43)	(1.22)			
Outside x n. events	-0.26***					
	(0.05)					
Outside x n. riots		-0.68*				
		(0.28)				
Outside x violence			-0.43***			
			(0.06)			
Distance				-9.14*	-11.57^{+}	-7.82**
				(3.63)	(6.22)	(2.48)
Distance x n. events				0.05		
				(0.04)		
Distance x riots					0.15^{+}	
					(0.09)	
Distance x violence						0.03^{+}
						(0.01)
R-squared	0.41	0.01	0.58	0.24	0.00	0.39
N	699	699	699	699	699	699

$^{+}p < 0.10$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

TABLE 8a.

SAC model: contiguity and residual contiguity

	(1) N. Events	(2) N. Events	(3) Riots	(4) Riots	(5) Violence	(6) Violence
Contiguity						
Lambda	0.70***	0.05	0.29***	0.13	0.79***	0.05
	(0.07)	(0.05)	(0.08)	(0.08)	(0.07)	(0.06)
Rho	-0.49***	0.26***	0.24*	0.33***	-0.64***	0.28***
	(0.13)	(0.06)	(0.11)	(0.05)	(0.13)	(0.08)
Capital	1.65	0.43	0.92	0.32	-2.73	-1.72
	(14.49)	(7.25)	(3.14)	(1.52)	(12.43)	(5.92)
Access	2.18	1.74	1.20	0.68	0.61	0.71
	(1.82)	(1.11)	(1.17)	(0.58)	(0.88)	(0.69)
Outside	-1.46	-2.10	-1.28	-0.23	-1.58	-2.09^{+}
	(2.32)	(1.36)	(1.49)	(0.72)	(1.66)	(1.23)
Distance	-0.02	-0.02*	0.00	0.00	-0.01	-0.02**
	(0.02)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)
Lagged dependent		0.68***		0.65***		0.67***
		(0.04)		(0.08)		(0.06)
R-squared	0.02	0.53	0.08	0.42	0.00	0.56
N	764	764	764	764	764	764
Resid. Contiguity						
Lambda	0.55***	0.05^{+}	0.15***	0.08	0.64***	0.05^{+}
	(0.10)	(0.03)	(0.03)	(0.06)	(0.07)	(0.03)
Rho	-0.37**	0.13***	0.14***	0.19**	-0.49***	0.12**
	(0.14)	(0.04)	(0.03)	(0.06)	(0.10)	(0.04)
Capital	1.39	-2.65	3.01	0.52	-2.95	-5.55
	(15.24)	(7.61)	(2.07)	(0.94)	(12.68)	(6.74)
Access	1.57	1.11	0.60	0.45	0.25	0.19
	(2.11)	(1.21)	(1.50)	(0.56)	(0.89)	(0.67)
Outside	-0.07	-0.95	-1.35	-0.21	-0.00	-0.97
	(2.54)	(1.73)	(1.79)	(0.76)	(1.52)	(1.38)
Distance	-0.02	-0.03*	0.00	0.00	-0.02	-0.02**
	(0.02)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)
Lagged dependent		0.70***		0.83***		0.69***
		(0.04)		(0.07)		(0.06)
R-squared	0.02	0.57	0.09	0.59	0.00	0.57
N	592	592	592	592	592	592

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

TABLE 8b.

SAC model: ethnicity and residual ethnicity

	(1) N. Events	(2) N. Events	(3) Riots	(4) Riots	(5) Violence	(6) Violence
Ethnicity						
Lambda	0.75***	-0.00	0.40***	0.19	0.74***	0.23
	(0.07)	(0.19)	(0.07)	(0.14)	(0.08)	(0.14)
Rho	-0.61***	0.33^{+}	0.33*	0.42***	-0.58**	0.04
	(0.16)	(0.19)	(0.14)	(0.08)	(0.18)	(0.19)
Capital	-3.12	0.59	1.23	0.60	-11.07	-2.89
	(16.95)	(6.99)	(3.27)	(1.53)	(14.94)	(6.66)
Access	2.97	1.37	0.47	0.12	1.63	0.86
	(1.93)	(1.26)	(1.17)	(0.58)	(1.14)	(0.80)
Outside	-0.61	-1.06	-0.13	0.61	-1.14	-2.11
	(2.51)	(1.38)	(0.66)	(0.39)	(1.88)	(1.48)
Distance	-0.02	-0.02*	-0.00	-0.00	-0.02	-0.02*
	(0.02)	(0.01)	(0.00)	(0.00)	(0.02)	(0.01)
Lagged dependent		0.65***		0.64***		0.61***
		(0.03)		(0.08)		(0.05)
R-squared	0.03	0.51	0.09	0.44	0.00	0.49
N	699	699	699	699	699	699
Resid. Ethnicity						
Lambda	0.64***	0.15*	0.59***	0.01	0.65***	0.26**
	(0.07)	(0.06)	(0.14)	(0.23)	(0.07)	(0.08)
Rho	-0.48***	0.07	-0.11	0.46**	-0.55***	-0.13
	(0.12)	(0.08)	(0.26)	(0.17)	(0.12)	(0.11)
Capital	-6.85	-5.37	0.88	0.11	-11.13	-7.53
	(18.24)	(8.44)	(3.01)	(1.34)	(15.20)	(8.09)
Access	2.93	1.54	0.65	0.13	1.59	0.90
	(2.03)	(1.27)	(1.19)	(0.64)	(1.15)	(0.83)
Outside	-1.85	-1.99	-0.59	0.65	-0.49	-2.64
	(2.61)	(1.49)	(0.73)	(0.45)	(2.00)	(1.61)
Distance	-0.03^{+}	-0.02*	0.00	-0.00	-0.03^{+}	-0.02*
	(0.02)	(0.01)	(0.00)	(0.00)	(0.02)	(0.01)
Lagged dependent		0.64***		0.66***		0.59***
		(0.03)		(0.09)		(0.04)
R-squared	0.03	0.49	0.08	0.44	0.00	0.44
N	694	694	694	694	694	694

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

TABLE 9a.

SAR model: internal and external contiguity

	(1) N. Events	(2) N. Events	(3) Riots	(4) Riots	(5) Violent	(6) Violent
Internal Contiguity						
Lambda	0.45***	0.27***	0.44***	0.33***	0.45***	0.27***
	(0.04)	(0.04)	(0.07)	(0.05)	(0.05)	(0.04)
Capital	8.24	6.52	2.50	3.43**	-6.21	-0.73
	(18.82)	(8.51)	(3.10)	(1.05)	(16.19)	(7.25)
Access	3.31+	2.13+	1.16	0.79	1.48	0.93
	(1.92)	(1.14)	(1.15)	(0.59)	(1.02)	(0.70)
Outside	-1.35	-0.39	-0.26	0.27	-0.79	-0.77
	(1.60)	(1.27)	(0.77)	(0.49)	(1.01)	(1.25)
Distance	-0.01	-0.01	0.00	0.00	-0.02	-0.01
	(0.02)	(0.01)	(0.00)	(0.00)	(0.02)	(0.01)
Lagged dependent		0.62***		0.61***		0.61***
		(0.04)		(0.07)		(0.07)
R-squared	0.01	0.59	0.06	0.54	0.00	0.60
N	762	762	762	762	762	762
External Contiguity						
Lambda	-0.00	0.00	-0.02	0.00	-0.00	0.01
	(0.01)	(0.00)	(0.02)	(0.01)	(0.01)	(0.01)
Capital	-0.58	0.40	8.21***	1.91*	-13.48	-5.63
	(24.80)	(9.89)	(1.63)	(0.79)	(21.87)	(9.26)
Access	0.06	0.69	-1.29	-0.41	0.56	0.72
	(2.57)	(1.44)	(1.36)	(0.44)	(1.47)	(0.84)
Outside	2.47	-0.67	-1.06	-0.81	3.10**	0.39
	(1.77)	(2.05)	(1.42)	(1.08)	(1.14)	(1.01)
Distance	-0.05	-0.03*	0.00	0.00	-0.04	-0.03*
	(0.03)	(0.01)	(0.00)	(0.00)	(0.03)	(0.01)
Lagged dependent		0.67***		0.91***		0.64***
		(0.03)		(0.09)		(0.05)
R-squared	0.02	0.51	0.02	0.61	0.02	0.43
N	447	447	447	447	447	447

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

TABLE 9b.

SAR model: internal and external residual contiguity

	(1) N. Event	(2) N. Event	(3) Riots	(4) Riots	(5) Violence	(6) Violence
Internal R. Contig.						
Lambda	0.38***	0.23***	0.26***	0.16**	0.34***	0.20***
	(0.05)	(0.05)	(0.07)	(0.05)	(0.04)	(0.03)
Capital	-1.79	1.36	3.39^{+}	0.83	-12.11	-3.71
	(21.69)	(10.07)	(2.00)	(0.89)	(19.08)	(8.76)
Access	2.21	1.61	0.83	0.61	0.70	0.43
	(2.32)	(1.33)	(1.60)	(0.61)	(1.02)	(0.68)
Outside	0.95	1.57	-0.64	0.41	0.80	0.98
	(2.23)	(1.15)	(1.51)	(0.66)	(1.76)	(0.75)
Distance	-0.02	-0.01	0.00	0.00	-0.02	-0.01
	(0.03)	(0.01)	(0.00)	(0.00)	(0.02)	(0.01)
Lagged dependent		0.63***		0.81***		0.62***
		(0.05)		(0.08)		(0.07)
R-squared	0.03	0.60	0.08	0.59	0.00	0.61
N	536	536	536	536	536	536
External R. Contig.						
Lambda	0.00	0.00	-0.01	0.01	0.01	0.01
	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.01)
Capital	-9.68	-3.47	7.50**	2.03*	-20.90	-9.02
	(27.86)	(11.27)	(2.30)	(0.97)	(24.48)	(10.57)
Access	-1.44	0.19	-2.37	-0.84	-0.10	0.63
	(3.46)	(2.00)	(2.00)	(0.62)	(1.76)	(1.05)
Outside	3.46	-1.29	-0.89	-0.81	3.82*	0.02
	(2.33)	(2.76)	(1.42)	(1.21)	(1.49)	(1.43)
Distance	-0.05	-0.03*	0.00	0.00	-0.04	-0.03*
	(0.03)	(0.01)	(0.00)	(0.00)	(0.03)	(0.01)
Lagged dependent		0.68***		0.90***		0.64***
		(0.03)		(0.10)		(0.05)
R-squared	0.02	0.49	0.06	0.61	0.01	0.42
N	307	307	307	307	307	307

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

TABLE 9c.

SAR model: internal and external ethnicity

	(1) N. Events	(2) N. Events	(3) Riots	(4) Riots	(5) Violent	(6) Violent
Internal Ethnicity						
Lambda	0.41^{***}	0.24^{***}	0.53^{***}	0.39^{***}	0.40^{***}	0.25^{***}
	(0.05)	(0.05)	(0.10)	(0.07)	(0.04)	(0.04)
Capital	16.71	12.87	3.11	2.00	6.15	8.16
	(20.18)	(9.69)	(3.66)	(1.99)	(16.59)	(7.66)
Access	2.85	1.57	0.52	0.34	1.60	0.76
	(2.27)	(1.39)	(1.21)	(0.60)	(1.27)	(0.89)
Outside	-1.31	-0.38	0.36	0.86^{**}	-1.70	-1.37
	(1.72)	(1.44)	(0.70)	(0.32)	(1.06)	(1.36)
Distance	-0.02	-0.01	0.00	0.00	-0.02	-0.01
	(0.03)	(0.01)	(0.00)	(0.00)	(0.02)	(0.01)
Lagged dependent		0.61^{***}		0.56^{***}		0.60^{***}
		(0.03)		(0.05)		(0.05)
R-squared	0.04	0.48	0.07	0.39	0.00	0.47
N	628	628	628	628	628	628
External Ethnicity						
Lambda	0.02	0.01	-0.05^{***}	-0.02^{***}	0.05^{+}	0.03^{*}
	(0.02)	(0.01)	(0.01)	(0.00)	(0.02)	(0.01)
Capital	-10.23	-3.31	4.64^{*}	0.84	-19.17	-8.02
	(20.04)	(7.97)	(1.97)	(0.81)	(17.45)	(7.64)
Access	2.44	1.30	0.10	-0.02	1.83	1.01
	(2.63)	(1.46)	(1.58)	(0.47)	(1.46)	(0.96)
Outside	-1.32	-0.50	-0.65	0.23	-1.15	-0.98
	(1.88)	(1.38)	(0.75)	(0.35)	(1.48)	(1.35)
Distance	-0.05^{*}	-0.03^{**}	0.00	-0.00	-0.05^{*}	-0.03^{**}
	(0.03)	(0.01)	(0.00)	(0.00)	(0.02)	(0.01)
Lagged dependent		0.67^{***}		0.91^{***}		0.62^{***}
		(0.03)		(0.04)		(0.04)
R-squared	0.01	0.52	0.10	0.73	0.00	0.43
N	611	611	611	611	611	611

$^{+}p < 0.10$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

TABLE 9d.

SAR model: internal and external residual ethnicity

	(1) N. Events	(2) N. Events	(3) Riots	(4) Riots	(5) Violence	(6) Violence
Internal R. Ethnic.						
Lambda	0.34***	0.22***	0.42***	0.29***	0.25***	0.16***
	(0.04)	(0.03)	(0.09)	(0.06)	(0.03)	(0.02)
Capital	7.90	8.01	-0.08	-0.10	-2.21	1.98
	(24.76)	(12.01)	(3.81)	(1.94)	(19.89)	(9.43)
Access	1.92	0.59	0.58	0.32	1.35	0.40
	(2.22)	(1.24)	(1.31)	(0.66)	(1.38)	(0.90)
Outside	-2.36	-0.86	-0.20	0.64^{+}	-1.83^{+}	-1.46
	(1.78)	(1.74)	(0.89)	(0.36)	(1.01)	(1.46)
Distance	-0.02	-0.01	0.00	-0.00	-0.02	-0.02
	(0.03)	(0.02)	(0.00)	(0.00)	(0.03)	(0.01)
Lagged dependent		0.60***		0.59***		0.59***
		(0.03)		(0.06)		(0.05)
R-squared	0.05	0.46	0.10	0.43	0.00	0.44
N	594	594	594	594	594	594
External R. Ethnic.						
Lambda	0.02	0.01	-0.04***	-0.01**	0.04^{+}	0.03*
	(0.02)	(0.01)	(0.01)	(0.00)	(0.02)	(0.01)
Capital	-10.38	-3.47	4.61*	0.84	-19.54	-8.39
	(20.09)	(8.08)	(1.96)	(0.80)	(17.61)	(7.85)
Access	2.51	1.26	0.08	-0.03	1.91	1.00
	(2.57)	(1.43)	(1.55)	(0.46)	(1.43)	(0.94)
Outside	-1.29	-0.48	-0.61	0.24	-1.10	-0.94
	(1.87)	(1.38)	(0.75)	(0.35)	(1.45)	(1.34)
Distance	-0.05*	-0.03**	0.00	-0.00	-0.05*	-0.03**
	(0.03)	(0.01)	(0.00)	(0.00)	(0.02)	(0.01)
Lagged dependent		0.67***		0.91***		0.61***
		(0.03)		(0.04)		(0.04)
R-squared	0.01	0.52	0.10	0.73	0.00	0.42
N	622	622	622	622	622	622

$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Chapter 3:
Ethnic Conflict as a Contagious Process

Carrie Huffaker
Università Commmerciale Luigi Bocconi

October 6, 2017

1 INTRODUCTION

Syria, Libya and Yemen: Three states that prior to the Arab Spring protest movement, enjoyed relative stability, if not expansive political freedoms. But conflict can sweep across borders like an epidemic, infecting one society after another until entire regions are aflame. The first wave of Arab spring protests led to peaceful regime change in Tunisia and Egypt. But in second-wave states like Syria, nonviolent resistance movements have given way to civil war, a war that has merged with ongoing sectarian strife in Iraq and with the violent Kurdish secessionist movement in Turkey.[1] Much of the Middle East is now aflame, generating the most severe refugee crisis since the Second World War. Collective social identities have proven crucial to the entire process: To the onset of conflict, and to its propagation.

In the past, to be Alawite or Christian or Sunni was perhaps not as important as the Syrian identity, but as conflict has crossed borders, sectarian and ethnic affiliations have acquired critical importance. Far-reaching national identities that once promised to encompass a plurality of peoples are under threat. Islamic State—a rapidly-growing Salafist movement dedicated to rebuilding the transnational caliphates that ruled the Middle East in the centuries following the Prophet Muhammad's death—has taken on the elimination of nationhood as a primary objective. In the territories they have conquered, the Iraqi-Syrian border no longer exists. It has been erased, physically and in the collective memory of the Iraqis and Syrians who support Islamic State. Only one identity can exist, in this conception of the world: That of the *ummah*, of the community of adherents to Sunni Islam.[2] War across the region and against a wide variety of other groups, from Yazidis to Shi'ites to the Druze, has been the inevitable result.

Epidemics of violence are not unique to the Arab world. In sub-Saharan Africa, war has spread from Rwanda to Congo to Uganda to Zimbabwe and elsewhere, all sparked by the mass murder of Tutsis in a single country in 1994 (Stearns, 2012; Meredith, 2007). In this case, as with the Arab Spring, ties between ethnic groups ruled by different sovereign regimes have emerged supreme as a channel for the international proliferation of civil war.

Ethnic and sectarian identities are not the only cause, nor even the primary one, of most

[1]See Lynch (2014) and Brownlee *et al.* (2015) for academic considerations of the Arab Spring and its aftermath, as well as of the role played by transnational ethnic and religious identities.

[2]See Wood (2015) and the 2014 VICE News documentary *The Islamic State*.

wars; but they are very often present when bloodshed crosses borders. In this paper, I propose a theoretical framework for explaining conflict as a contagious process that hinges upon transnational ethnic ties. I explore how resource wealth, as well as the institutions guiding the distribution of economic and political power among different groups, both shape the likelihood that conflict emerges in a single, isolated state. The situation changes when an epidemic of violence emerges and begins to disseminate from country to country. I argue that under threat of contagion, the key variable separating societies into states of peace and states of war is β, a stochastic parameter that governs spillovers from ethnic conflict abroad. This parameter represents the strength of transnational ethnic ties.

I present two models. Each one focuses upon a different aspect of the problem and thus contributes to our understanding of the propagation of conflict in a unique way. The first model, based upon Besley and Persson (2011), relates transmission processes to the strength of institutional constraints binding the behavior of the ruling regime. In the second model, based upon Morelli and Rohner (2015), I identify risk of contagion with the shock to the *de facto* power of marginalized ethnic groups that can materialize with the outbreak of conflict abroad.

I hypothesize that spillovers impact conflict risk primarily by shifting the incentives of minority groups belonging to the opposition. First, minority groups may no longer accept the status quo system, developing new preferences over redistribution. Second, they may build up larger fighting forces. Minority groups cannot commit to refrain from applying their new strength to challenge the government. The regime, in turn, cannot credibly commit to a strategy of peaceful appeasement. The two groups are unable to bargain, and violence spreads. Countries lacking all the factors usually associated with conflict risk are plunged into warfare due to transnational ethnic ties to violent neighbors.

The organization of the paper is as follows. In section 2, I review how different literatures explain the onset and spread of conflict. I discuss four potential channels for conflict contagion found in the literature. I explain how my own framework contributes to the literature in the following section. In section 4, I present the two models and derive a number of interesting theoretical results. The paper closes with a discussion and some concluding remarks.

2 LITERATURE REVIEW

In this paper, I present a new theoretical model of conflict contagion through transnational ethnic ties. This model is deeply embedded in the existing body of work on ethnic conflict and civil war.[3] Therefore, this literature review briefly reviews the main trends in conflict literature since the 1990s, explaining how my framework relates to and builds upon each one. I begin by discussing the literature of national and local factors that can explain the onset of violent conflict, before moving on to transnational factors and finally, a networks approach. The scope of analysis thus becomes progressively larger. Proceeding in this way is important to our understanding of ethnic conflict not merely as a national or local process, but as a phenomenon that can tear apart entire regions, reinforcing itself through channels of contagion. I conclude the review by introducing my model and discussing, in section 3, its unique contribution to the economic literature on the causes of war.

2.1 NATIONAL AND SUBNATIONAL FACTORS

Before turning to contagion and network effects, it is useful to consider explanations for the onset of civil war that center upon the characteristics of individual nation-states. These factors determine the baseline risk of conflict, without accounting for transnational factors. They are important in explaining the emergence of conflict in "patient zero," the country where the epidemic originates. Moreover, countries structurally susceptible to strife are far more vulnerable to spillover processes.

Most country-level models in use today adopt as their point of departure the conflict-as-crime materialist literature of the 1990s. Extensions of this literature place special emphasis upon the roles of resource scarcity, resource wealth and political institutions. Later work also considers subnational and transnational factors, the focus of my paper. Common to all approaches discussed in the literature review is the fundamental idea that any factor that either enhances the incentives of potential insurgents to rebel, or weakens the constraints placed upon them, will strengthen their position vis-à-vis the government and heighten the likelihood of conflict.

[3] I employ the terms *ethnic conflict* and *civil war* somewhat interchangeably, as it is difficult to judge where one ends and the other begins. Moreover, my model is based on the conceit that the two cannot be cleanly disentangled. I delve into a deeper consideration of the meaning of *ethnicity* at the end of Section 3.

2.1.1 MATERIALIST EXPLANATIONS FOR CIVIL WAR

The treatment of civil conflict as a society-large competition over the spoils from criminal activity emerged in the late 1990s in the work of Collier and Hoeffler (1998, 2002) and Fearon and Laitin (2003), among others. Much subsequent work in the field essentially follows from these papers.[4] In U.S. criminal law, the establishment of means (available resources), opportunity (commitment of means towards "unforeseen gains, but containing an element of risk") and motive ("the inducement, reason, or willful desire and purpose behind the commission of an offense") is necessary to the determination of a guilty verdict.[5] The insurgency-as-crime authors focus upon the first two elements, arguing that the motivations for rebellion—ethnic and politic grievances—are essentially universal and thence unable to explain the pattern of civil conflict that we observe. Insurgencies instead occur wherever they are feasible, that is when means and opportunity create a favorable "opportunity structure" for resistance and rebellion.

The cross-national, large-N empirical studies from this period uncover patterns of correlation between conflict onset and economic indicators such as low p.c. income and dependence upon primary commodity exports, demographic and social factors like population and a large foreign diaspora, and geographic determinants like mountainous terrain (*Ibid*). Recent conflict and political instability are powerful predictors of renewed outbreak. Evidence for motive indicators such as ethnic dominance, fractionalization and polarization is contradictory and weak, lending credence to the idea that they are unimportant.[6]

The early body of materialist research fell short in a number of ways. Focusing upon large-scale national conflict ignores the localized, low-intensity nature of most civil unrest. Country-level indicators of ethnic diversity, in particular, fail to reflect the actual balance

[4] The pioneering work of anthropologist Jared Diamond (1997), who models economic growth as a deterministic function of geography and ecology, is arguably a precursor to the materialist civil war literature in economics.

[5] See the online edition of *Black's Law Dictionary* at http://thelawdictionary.org/motive/.

[6] Collier and Hoeffler (2002) report a positive correlation between conflict risk and ethnic dominance, and a negative one with ethnic fractionalization. Fearon and Laitin and Miguel *et al.* (2004) fail to find a robust relationship for any grievance indicator. Hegre and Sambanis (2006) instead find that ethnic diversity is a significant predictor of lower-intensity conflict in their review of the literature. More recent work suggests a positive correlation between measures of ethnolinguistic fractionalization (Harari and La Ferrara, 2013) and polarization (Esteban *et al.*, 2015; Fosberg, 2008) and the likelihood of conflict onset.

of power at the local level (Cederman *et al.*, 2010). Moreover, such fundamental "causes" of war as poverty and resource dependence are as likely to be the endogenous outcome of conflict (resources #1, resources #3).

A number of new schools of thought about the causes of civil war have developed over the last decade, always evolving from or in reaction to the early materialist work. They link incidence of conflict to a scarcity of renewable resources (section 2.1.2), to nonrenewable resource wealth (2.1.3), and to the institutional framework (2.1.4). The first two schools investigate how different types of natural resources influence the opportunity structure. The influence of resources is thought to operate through several mechanisms. Relative scarcity decreases the opportunity cost of fighting by reducing wages, while relative abundance increases it. Shocks to resource wealth also change the size of the "prize" to be fought over, as well as the probability of winning a conflict (depending upon the distribution of power and property rights across groups). The third school analyzes how institutional constraints influence the redistribution of resources and power between different groups in a country. Several authors have already begun to introduce transnational factors and contagion processes into these three bodies of literature, as I will discuss below.

2.1.2 RESOURCE SCARCITY

Addressing the endogeneity of wealth discussed above, Miguel *et al.* (2004) introduced the now-popular methodology of instrumenting economic growth by rainfall. Diamond (2005) has linked an inability to adapt to environmental change to the collapse of a number of civilizations across history, and increasing concern about global warming has inspired scientists across many fields to study the relationship between climate variability and modern armed conflict (Burke *et al.*, 2009; Buhaug, 2010). Recent research has taken advantage of the increased availability of disaggregated data to model local-level processes. Much of the work uses rainfall data from satellites and gauges to map how climatic volatility influences local livelihoods and hence, the risk that regions will be struck by violence. However, this literature is only capable of explaining the onset of conflict conflict in sub-Saharan Africa: The one region of the world where a majority of livelihoods still depend upon rain-fed agriculture. Environmental instability plays a much weaker role elsewhere (Miguel; Gleditsch, 2012).

While some work finds negative deviations in rainfall to be correlated with heightened conflict risk (Miguel; Harari and La Ferrara, 2013), other authors argue that droughts can induce cooperation between competing ethnic groups. Rainfall abundance may pose a greater risk (Gleditsch, 2012). There is some evidence that flooding is even more dangerous than drought, as it can disrupt livelihoods in more extreme ways and generate destructive patterns of scarcity (*Ibid*). Other work finds extreme rainfall patterns to be robustly associated with drops in growth, but not in conflict. Indeed, no link between conflict risk and scarcity of any kind of renewable resource– as characterized by drought, soil degradation, deforestation or population density—remains robust to alternate specifications and estimation strategies (Buhaig *et al.*, 2008).

The Harari and La Ferrara model also analyzes spatial dependence in conflict induced by negative shocks to agriculture. Working on a very disaggregated scale, they find that spillovers across time and space are large and significant.[7] Their work is therefore an important empirical link between the resource scarcity and contagion literatures.

2.1.3 RESOURCE WEALTH

Unlike scarce renewables, a number of nonrenewable resources appear robustly linked to conflict onset: Petroleum, natural gas and alluvial diamonds (Aragon *et al.*, 2015; Ross, 2013). These "point-source," extractable resources are geographically concentrated and easy to appropriate. Their presence incentivizes rapacity and rent-seeking on the part of both rebels and the state. Shocks to natural resource wealth, such as the discovery of new oilfields, can drastically destabilize a society. Furthermore, not only geographic location, but also the type of resource sector in which wealth is concentrated, appear to matter for predation and violence. Exogenous price shocks to capital-intensive sectors such as oil tend to increase strife, while shocks to labor-intensive sectors such as coffee reduce it (Dal Bó and Dal Bó, 2011; Dube and Vargas, 2013).

The distribution of point-source resource across the geographic territory of a state can predict border disputes, ethnic conflict and even genocide: Phenomena almost universally seen as motive-driven. Ethnic groups concentrated in oil- and diamond-rich territories appear to be targets for strategic violence (Morelli and Rohner, 2015). Caselli *et al.*

[7] The presence of conflict in all of a local cell's neighbors increases the probability of conflict in the surrounded cell by 26 percentage points.

(2015) study the onset of cross-border wars in relation to the location of oilfields. Such models suggest how the geographic distribution of resources could also influence contagion of ethnic conflict across borders.

The wealth-conflict interaction appears to be nonmonotonic, with the poorest and richest countries at greatest risk. Moreover, the intermediating role of the political and legal system is often vital. The resource curse appears only to infect those countries with ex-ante weak, "grabber-friendly" institutions (Aragon *et al.*; van Der Ploeg, 2010). However, the direction of causality is ill-defined, as oil windfalls increase corruption and decrease accountability, leading to the deterioration of institutional quality (*Ibid*). Because the phenomena of resource wealth and institutional evolution are so closely linked, the two literatures are intertwined. The impact of resources upon conflict is usually analyzed through the lens of political and economic institutions.

2.1.4 INSTITUTIONS

Institutions are arguably the fundamental determinant of long-run economic growth (Robinson *et al.*, 2005; Robinson *et al.*, 2003), and thus indirectly influence the risk of conflict via the means-opportunity mechanisms. They inform a country's opportunity structure. A number of authors also posit a direct link between institutions and conflict. Like resource wealth, institutions may be correlated with conflict onset in a nonmonotonic way; empirical evidence suggests that the more extreme cases—established democracies and strict autocracies—are less at risk of conflict (Hegre *et al.*, 2001).

Democratic regimes are severely constrained in their behavior by existing institutions, yet this allows them to channel grievances into the legislative process and to credibly commit to change. Strict autocracies are unconstrained because institutions are weak and subverted to the personality of the ruler, enabling the autocrat to brutally suppress any hint of unrest; however, they are unable to make any sort of credible commitment to reform (Acemoglu and Robinson, 2009). Institutional strength can thus be understood as the degree to which institutions constrain the behavior of the ruling regime and render promises enforceable. Strong institutions incorporate a broad cross-section of society into the political process and redistribute resources in what is perceived to be an equitable way. Weak ones allow regimes to appropriate the national territory and everything in it as their personal fiefdom.

Democracies have strong institutions and a great commitment power, while autocracies cannot make any promises that their populations will ascribe to, but paradoxically each is far more robust to unrest than are systems characterized by mixed institutions, or in the process of transitioning towards democracy.[8] New democracies may be at greatest risk of violence in the form of ethnic cleansing and even genocide, as regimes seek to consolidate their power (Mann, 2005).

Of all systems of political governance, established democracies appear to be the least prone to violence. This is largely due to their ability to channel grievances. Besley and Persson (2011) find that lack of consensuality and inclusiveness in institutional environments drive conflict. Large ethnic groups, excluded from or underrepresented in state power—as well as those recently driven from power—are far more likely to challenge the ruling regime in a violent way (Buhaug and Gleditsch, 2008; Cederman *et al.*, 2009): "There is an unequivocal (negative) relationship between degree of access to state power, and likelihood of armed rebellion" (Cederman *et al.*, 2010). Strong institutions are thus the greatest safeguard against violence. They are also essential to understanding how a state will respond to the risk of contagion from foreign conflicts.

The factors explored by the resource scarcity, resource wealth and institutions literature are crucial to the formation of a state's domestic strategic environment. They determine the incentives and ability of groups to credibly commit to post-conflict strategies and ultimately, whether bargaining can be successfully concluded. When opposing forces fail to reach such agreements, violence breaks out. In exploring resources and institutions we are therefore identifying the key factors that drive an isolated country to civil war. Yet no country in the world is isolated.[9] By introducing transnational forces, I shift focus from internal forces to the larger strategic environment in which the single nation-

[8] Ancient Rome is an example of how mixed and transitioning systems of governance can give way to violence. As Rome's empire grew, the decaying institutions of the Republic (designed for a contained city-state) were unable to cope with the requisites of ruling over vast and heterogeneous territories, opening the way to two centuries of crisis peopled by demagogues who sought to concentrate ever more power in their own persons (Marius, Catiline and Caesar, among others, embody this trend). Thus a long-established democracy devolved into the strict autocracy of Augustus. Under constant attack by barbarian incursions, the monarchy fell into a crisis which later emperors sought to counter through increasingly inclusive measures such as the extension of Roman citizenship to every free male in the empire (by Caracalla, in 212 C.E.). These attempts at democratization were ultimately unsuccessful, as the Western empire finally fell to German barbarians in 476 C.E. See Scarre (1995) and Tainter (1998).

[9] Even the "hermit kingdom" of North Korea depends heavily upon China, its closest ally, for survival and for intermediation with the rest of the world.

state is embedded. The ultimate aim of my paper is to investigate how such domestic factors interact with the strategic neighborhood surrounding each country. This will provide an exciting opportunity to integrate the resource and institutions literature with the transnational and networks literature I will discuss next.

2.2 TRANSNATIONAL FACTORS

Scholars have long noted a neighborhood effect, or clustering of civil conflict in specific regions such as the Great Lakes region of Africa and the Arab world (Sambanis, 2002). The question remains whether the observed neighborhood effect is due to the spatial clustering of country attributes conducive to strife, or to the actual spillover of conflict through processes of contagion. Much as the country-level analyses have borrowed the language and concepts of criminology, the more recent transnational literature builds upon the methods of epidemiology. Conflict is re-envisaged as an infectious disease. Researchers seek to identify such factors as the incidence (onset) rate, the relative risk according to different idiosyncratic characteristics, and what populations should be considered at risk of contagion and why.

There is robust evidence of a spillover effect, even once the geographic clustering of risk factors is accounted for (Buhaug and Gleditsch, 2008). Gleditsch (2007) argues that the marginal effects of changes in transnational and national factors are equally large: "The likelihood of civil war in an extremely unfavorable region would be several hundred percent higher than the risk of conflict in very favorable neighborhoods." A peaceful democracy surrounded by war-prone, undemocratic neighbors is in great danger (Hegre and Sambanis, 2006).

Four possible channels for the contagion effect are put forth in the literature. First is the direct provision of material and moral support for war between neighbors. War can flood entire regions with arms and combatants, potentially reducing the cost of insurgency elsewhere (Collier and Hoeffler, 1998). In addition, direct intervention can generate spillovers of all types across countries. Regimes (as well as non-ruling groups) often intervene more or less overtly in the conflicts of their neighbors, and aside from spreading violence and arms this can severely destabilize the situation at home: As occurred in Zimbabwe following Mugabe's deeply unpopular decision in 1998 to intervene in the Congo war (Meredith, 2007). In this case contagion occurred even though the two countries are separated by over 2000 kilometers.

Second, refugee flows bring conflict with them (Buhaug and Gleditsch, 2008). Large numbers of refugees worsen economic conditions in the host country by depressing wages and raising housing costs (Reuveny, 2007). By shifting the ethnic composition of the population, they threaten the existing balance of power and inspire reactionary movements (Lake and Rothchild, 1998). Moreover, their presence favors the exchange of resources, ideas and arms between insurgent groups. Perhaps most importantly of all, refugee camps are a fertile recruiting ground and sanctuary for rebel organizations (Salehyan and Gleditsch, 2006).

The third channel is the demonstration effect. Minority groups in a given country may observe unrest elsewhere and update their beliefs about the plausibility and desirability of challenging their own regime (Lake and Rothchild, 1998; Heydemann and Leenders, 2014). Transnational ethnic identities appear to be particular effective instruments of contagion. Forsberg (2008) argues that "if an ethnic group rebels in one state, it increases uncertainty among kin in neighboring states. With a change in the bargaining position and behavior of kin, group members across borders are confronted with a new decision calculus." Observing the rebellion of a shared ethnic minority against a neighboring regime, a government may anticipate the shifting of relative group strength on its own soil due to the demonstration effect, and choose to engage in preemptive repression (Danneman and Ritter, 2014).[10]

Finally, spillovers occur as unrest along ethnic lines contributes to the solidification of ethnic identities across the region (Gurr, 2000). Actors may rise up out of solidarity with ethnic kin located in other states (Salehyan and Gleditsch, 2006). Ethnicity is a crucial resource because it enables actors to overcome the greatest barrier to constructing a coherent protest movement: The collective action problem (Maves and Braithwaite, 2013). Ethnic identities are a powerful force for unification and mobilization; perhaps the most powerful of any, for they underlie the very notion of national sovereignty (Cederman *et al.*, 2010).

Scholars have long identified transnational ethnic ties as a source of contagion: Far more important to the spread of conflict, in fact, than geographic proximity (Maves and Braith-

[10]This employment of preemptive repression is similar to the mechanism underlying strategic mass killing in Esteban *et al.* (2015); both rely upon shocks to relative group strength.

waite). Each of the four channels—the direct provision of support, refugee flows, the demonstration effect, and the solidification of ethnicity—may operate partially or entirely via the action of transnational ethnic kin. These findings challenge the supremacy of means-opportunity explanations, opening the way for a new consideration of how grievances drive uprisings.[11] Yet the behavior of groups defined by ethnicity can still be analyzed from a rational point of view. A number of political economy models explore how institutional constraints shape contagion processes.

Abstracting from their surrounding neighborhoods, the countries most vulnerable to contagion are those with a high baseline risk of conflict: Poor, populous and polarized, with a past history of conflict (Buhaug and Gleditsch; Forsberg, 1998). These states lack the capacity to further repress or appease restive minorities in order to avoid conflict, and once transnational ethnic kin (TEK) rebel, the domestic balance of power shifts and they too are pushed into chaos (Danneman and Ritter). Cederman *et al.* (2013) find that the existence of very large or very small TEK groups, included in own their ruling coalitions, lowers the risk of contagion; while stateless TEK (such as non-Iraqi Kurds) always increase risk of conflict at home. Others find that alienation from power of TEK is unimportant.

Evidence is mixed regarding whether improved representation of domestic minorities is beneficial or dangerous. Some authors argue that any factor that enhances the ability of minorities to mobilize can potentially imperil the regime, once conflict breaks out abroad (Braithwaite, 2010; Cederman *et al.*, 2009). However, this is at odds with the findings of the institutional literature discussed above and may only pertain to strict autocracies with weak institutions.

The impact of rebellions abroad involving transnational kin of the domestic ruling coalition is also unclear. Little empirical research has addressed this aspect of the problem. Such rebellions might strengthen the domestic government's hold on power, by increasing its demographic weight through refugee flows and providing it with materials for war, or else threaten ruling authority through the demonstration effect or by exciting tensions between opposing groups. Future empirical work is needed to clarify this point. I focus only upon rebellions involving domestic minorities, excluded from the ruling coalition, for the purposes of this paper.

[11] Although Collier and Hoeffler (2002) identify large foreign diasporas as a risk factor, they argue that "greed considerably outperforms grievance" as a strategic consideration.

The weakness of the existing transnational literature lies largely in the fact that it is almost entirely empirical. The absence of a rigorous framework grounded in economic theory translates into a lack of clearly falsifiable hypotheses and predictions. Nor has this literature yet been into the larger body of theoretical work on the causes of civil war. In applying economic theory towards explaining the powerful ideas that emerge from the empirical studies of contagion discussed above, we take a necessary and important step towards enriching our understanding of the economic causes of ethnic conflict and civil war.

Figure 1 summarizes how domestic resource shocks and foreign ethnic conflict can interact to shape conflict onset. Together, these two types of shocks destabilize local livelihoods and the current balance of power between the governing coalition and ethnic groups belonging to the opposition.

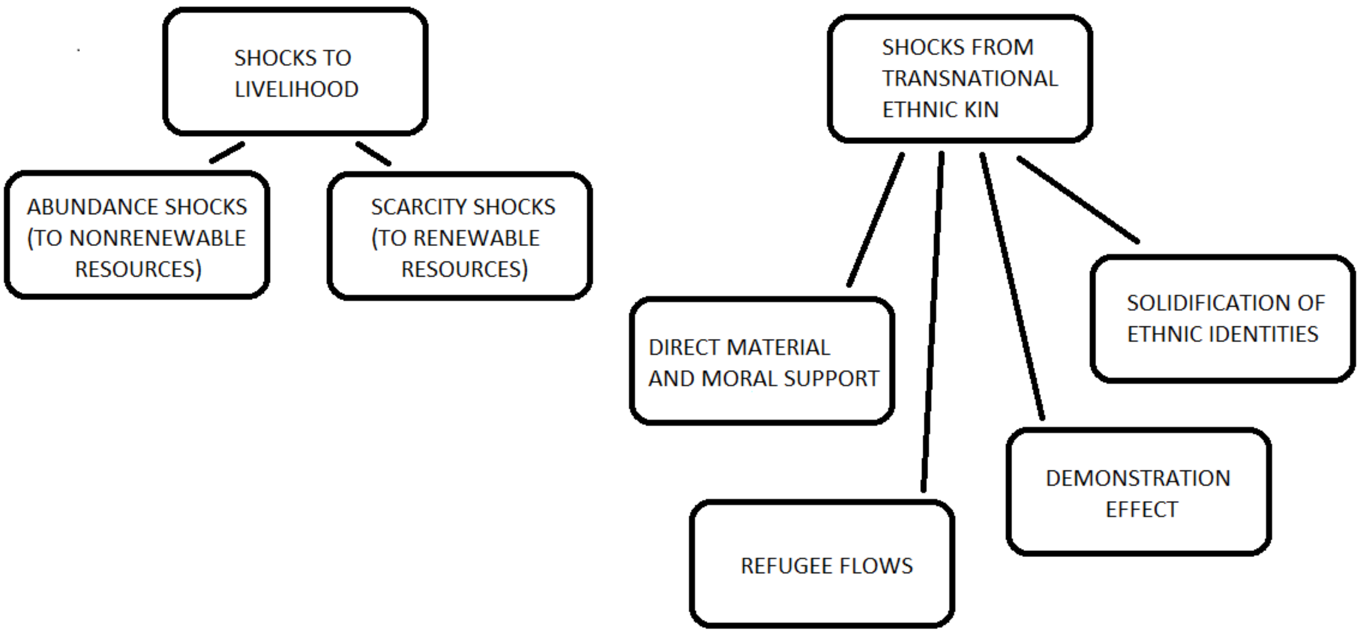

Figure 1: The impact of shocks upon domestic livelihoods and power

Figure 2 models how the final state of the world– peace or war– is jointly determined by the strategic response of state and of opposition to the transformation of the domestic environment. Political and economic institutions constrain both responses, thus directly impacting the likelihood of contagion. The result is a hierarchy of cause and effect, beginning from the initial shocks and ending with the two potential states of the world.

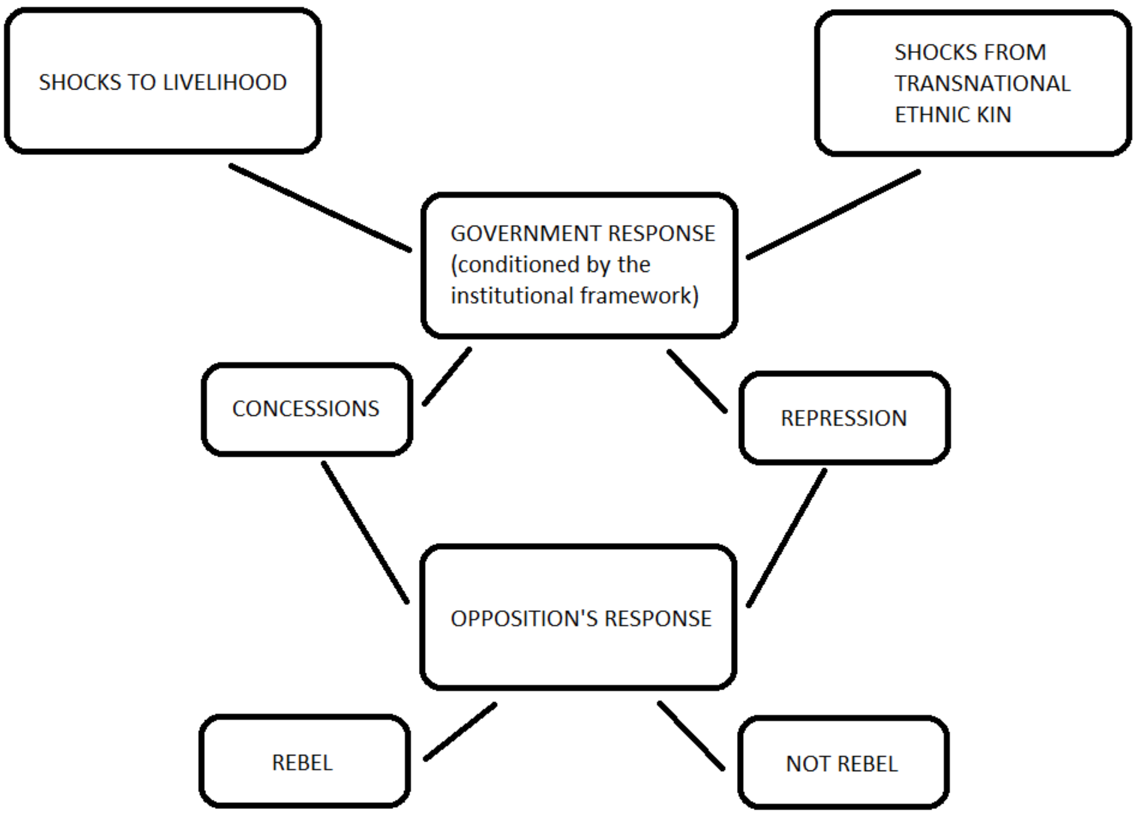

Figure 2: The transnational spillover model as a hierarchy of cause and effect

It is helpful, in analyzing the interaction between a state's domestic and international strategic environments, to reconceive the problem using tools from network theory. We can reconceptualize geographic regions as networks of interlinked nodes. In my case, the relevant links are transnational ethnic ties, which acquire significance once civil conflict erupts and begins to spread. In my models, I hypothesize that the fighting efforts of transnational ethnic kin, of domestic minorities, and of the government are all strategic complementarities, mutually reinforcing each other. Therefore in concluding the literature review, I briefly turn to how innovative new work in network theory models conflict.

2.3 NETWORKS

A number of recent papers explore how network architecture shapes behavior in strategic conflict games. Franke and Ozturk (2015) find that for several different types of networks, conflict intensity is increasing in the number of conflictual relationships and in the density of the network. Bramoullé *et al.* (2014), focusing upon games with strategic substitutes—in which case direct and indirect network effects work in opposite directions—find that equilibria depend upon the lowest eigenvalue of the network matrix. Konig *et al.* (2015) model the Second Congo War as a network of alliances and enmities, with strategic substitution in the fighting effort of allied groups. They find that for sufficiently large alliance externalities, the marginal benefit from fighting falls to zero and the network becomes completely peaceful. Conversely, a network of hostile links is plunged into total warfare.

The four channels of contagion suggest a different relationship between friendly transnational kinship groups. They suggest, in fact, that the decision to rebel of an ethnic group in one country will increase the relative strength, and therefore the incentive to fight, of kin elsewhere. The fighting effort of allied ethnic groups embedded should therefore be characterized by strategic complementarities, not substitution. Complementarities result from the fact that actors are fighting in separate arenas over different prizes, as in Franke and Ozturk. When an ally in the network increases its fighting effort, this does not generate contrasting incentives to both increase and decrease one's own effort. When allies (and enemies) invest more resources in war, the group under analysis will always increase its own effort.

Unlike in the case of substitutes, direct and indirect network effects are fully aligned under strategic complementarities. Thus no equilibrium exists when network externalities are large. The individual effort of agents creates a feedback cycle that soon spirals into infinity. Bounded action sets instead permit the existence of multiple equilibria. With linear best replies, the equilibrium is unique (Belhaj *et al.*, 2014). Equilibrium can be reached under strategic complements and large network effects because central agents will quickly arrive at their upper bound, and the interdependence between agents' effort levels is broken.

In my model, single dyads–linked together through ethnic ties– each fight over a unique prize. Moreover, the action sets of agents are circumscribed by resource constraints. Thus exploding effort levels are not a concern. This distinction highlights the fundamental difference between network and political economy models of conflict. In network models, there is no room for institutions or collective identities; there is no way to model the heterogeneous strategic incentives of different types of actors. Fighting effort is largely determined by network structure. Because my primary interest lies in how and why the channels of contagion operate, the tools of political economy are more useful.[12]

[12] Network models do enjoy the distinct advantage that the simultaneous choices of any number of agents can be modeled together. Thus there is a trade-off in adopting a political economy rather than networks approach.

Figure 3 shows how the theoretical framework that I will present in the next section can be embedded in a network-like setting. The network is composed of three different countries (1, 2 and 3) with two different marginalized ethnic groups, A and B. Group B is only present in country 2. Suppose conflict breaks out in country 2 between the ethnic group in power and the two marginalized minorities. In our model, ethnic identity is the salient principle in transnational ties. Conflict spreads to country 1 and 3 through group A— present in all three countries— but not through B. If conflict were to arise independently in country 1, this would not reflect contagion but merely the regional clustering of attributes conducive to ethnic conflict. Contagion occurs if the outbreak of war in 1 and 3 reflects the transnational ties of group A, and would not have occurred otherwise. By studying individual countries in isolation, we would never be able to make this distinction: Which is why a theoretical model of transnational contagion is needed.

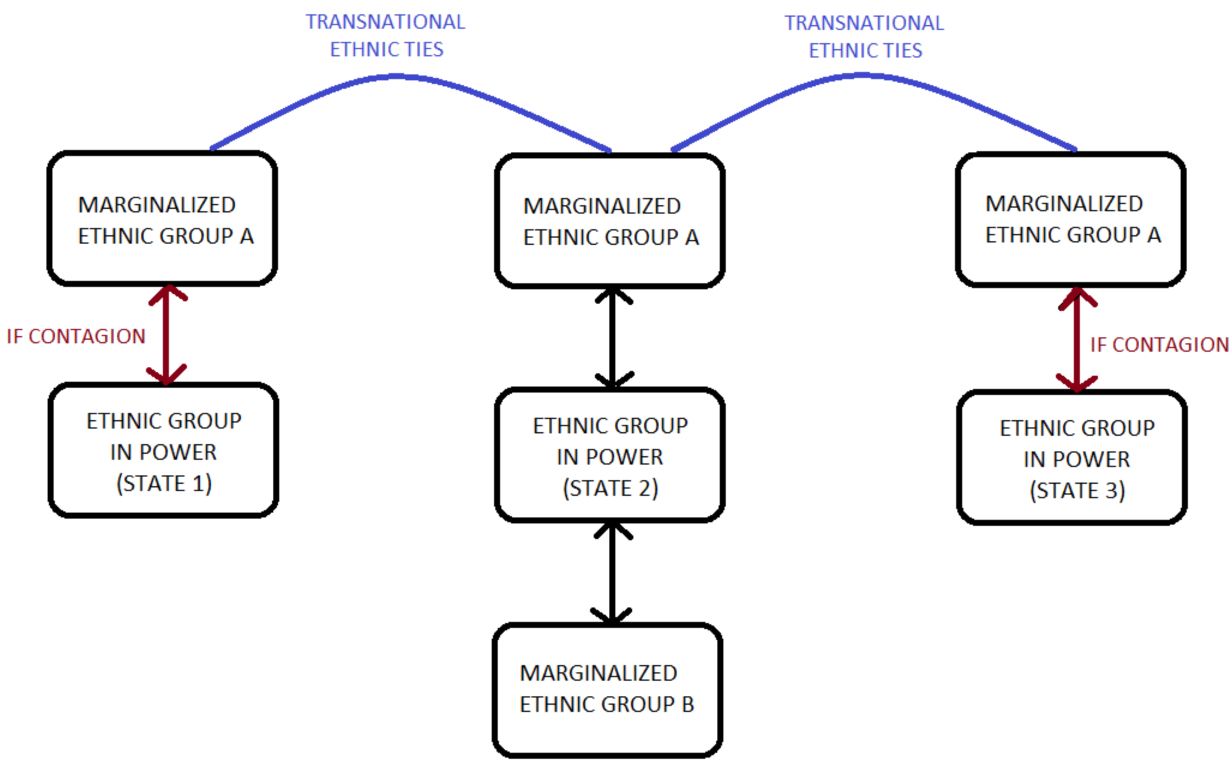

Figure 3: The transnational spillover model in graph form

3 CONTRIBUTION

I propose a new theoretical foundation for transnational spillovers, built upon two existing models from the resource wealth and institutions literature on the causes of civil war. The equilibrium of each model is characterized by three ordered states of the world, divided by threshold values of a key variable. The lower threshold conditions the outbreak of violence; the upper one determines whether a state is plunged into all-out civil war. In Besley and Persson (2011), the key variable is exogenous and stochastic: Resources available for redistribution between the two main coalitions in society. In Morelli and Rohner (2015), it is endogenous: The incumbent's decision rule for the division of resource wealth.

While resources and their allocation between the two coalitions continue to play an important role in the two models presented here, I introduce a new variable that crucially determines the threshold between war and peace: β, measuring the strength of spillovers from wars abroad involving transnational ethnic kin of the opposition.[13] I hypothesize that transnational spillovers impact the domestic bargaining processes in two key ways. First, they alter the preferences over redistribution of the opposition. Second, spillovers influence the fighting efforts of each group, and the probability of a successful insurgency.

The first spillover effect is pivotal to reshaping the behavior of the opposition. Observing allied kin groups rebelling against their own governments abroad, domestic minorities will no longer be content to accept the status quo system of redistribution. The term *redistribution* can be understood here in a double sense. The division of surplus rents accruing from diamonds, gold, petrol and other precious resources, as well as cash crops like coffee, sugar or coca, constitutes one form of redistribution; the division of political power, a second form. In keeping with the demonstration effect, foreign unrest can create a shift in perceptions of the domestic system, rendering it unacceptable in the eyes of minorities, even if these groups were unable or unwilling to challenge state power before. Groups belonging to the opposition's coalition will be driven to demand from the government a larger share of resource rents and more political power, particularly over their own territory.

Second, spillovers generate a shock to the relative strength of ethnic minorities vis-à-

[13] This variable can also be interpreted as measuring the strength of transnational ethnic ties, and should vary across time and space as the salience of different identities grows or weakens.

vis the governing coalition, creating a dissonance between *de facto* and *de jure* power in the country.[14] In keeping with the first, second and fourth channels of contagion, the opposition will have access to a larger supply of arms and recruits; a possible gain in demographic power; and a strengthened sense of identity, helping it to overcome the collective action problem. Ethnic warfare abroad essentially creates a shock to the *de facto* power of the opposition. In the first model, this new strength expresses itself as a direct shock to the size of the insurgency. In fact, both the opposition and incumbent increase their investment in military power. The probability of waging a successful insurgency rises in the second model. Due to the support of foreign kin, the opposition is better able to counter the advantage accruing to the ruling coalition due to control of the state apparatus.

Violence arises because the new distribution of power gives rise to commitment problems and therefore, the failure to reach an agreement that could assure peace between the two groups. Bargaining failure results because the minorities cannot credibly commit to avoid acting upon their changed incentives and their strengthened position to challenge government rule. The state—anticipating the enhanced strength of minorities in the opposition—cannot commit to avoid preemptively crushing the minority groups through brutal repression or all-out warfare. Each side therefore has the incentive to pour resources into its military capabilities, creating an escalation of military might. Therefore, fighting effort of ethnic kin abroad, of domestic ethnic minorities and of the government are all strategic complements. Violence can be thought of as a mutually reinforcing process. As minorities abroad build up their revolt movement, the marginal utility of violent rebellion increases for allied domestic minorities, as does the marginal utility of violent repression for the government.

These two changes induced by spillovers—to fighting strength, and to preferences over redistribution—create a new set of thresholds governing the three states of the world. The key innovation in this enrichment of the two models is that not only domestic factors, but the larger environment in which nation-states finds themselves embedded, inform the strategic behavior of government and opposition. Societies that would have remained

[14] Robinson *et al.* (2005) argue that *de jure* power (legal control of the state) is determined by political institutions, while *de facto* power is unofficial and shaped by economic institutions. *De facto* power determines a group's ability to overcome the collective action problem and is the sure path to consolidating *de jure* power. I hypothesize that ties to transnational ethnic kin constitute a second source of *de facto* power for marginalized ethnic groups.

peaceful and stable given their structural characteristics can find themselves pushed into civil war due to the influence of warring ethnic groups in neighboring states: Reflecting the finding of Gleditsch (2007) that even stable democracies are greatly at risk when encompassed by belligerent neighbors. The danger heightens even further when ethnic identities traverse international borders, offering a source of power and legitimacy that undercuts that of the state. Ties to transnational ethnic kin in warring states are therefore the crucial instrument of conflict contagion.

Because it is such a fundamental variable in our analysis, it is important to closely consider the definition and meaning of *ethnicity*. Cederman *et al.* (2010 why rebellion 6) define *nationalism* as the "political principle that the unit of governance and the nation should be congruent... ethnic likes ruling over ethnic likes." *Ethnic politics* is the struggle between different ethnically-defined polities over control of the state apparatus. Yet the question of ethnic identity is a sticky one. No group is wholly defined by a single metric. Kurds may practice the Sunni, Shi'a or Yazidi faith; Arabs may be Sunni, Shi'a, Druze, or Christian; Shiites may be Persian, Arab or Kurdish, among many other ethnicities. Differing collective identities are activated at different points in time.[15] It is useful to reconceive such conflicts with transnational dimensions as *wars over identity*, as Lynch (2014) does in his analysis of the Arab Spring, rather than restricting attention to a single dimension such as religion or language.

In this paper, I apply the term "ethnicity" to groups with a solid common identity and larger political ambitions. The *governing coalition* (or incumbent) therefore refers to the ethnic group controlling the state apparatus, and all other ethnic groups allied with it or enjoying its protection; while the remaining ethnic groups excluded from the government constitute the *opposition*. The opposition are also referred to as ethnic minorities, for they are marginalized from state power; although in many cases they are numerically superior to the dominant group.[16] It is shocks to this heterogeneous collection of non-governing

[15] The assimilated Jewish populations of Western Europe often considered themselves to be German, Austrian or French first, and Jewish only second, prior to the persecutions of the Nazi period. In the course of the Holocaust, this secondary identity acquired primary importance.

In the case of Eastern Europe, millions of ethnic Germans (a legacy of the Hapsburg Empire) were expelled at the end of the war because these nations no longer viewed them as fellow countrymen but as the enemy. Hundreds of thousands were killed. Most fled to Germany or Austria. See Gilbert (1985), King (2002) and the 1997 documentary *The Long Way Home*.

Both of these examples demonstrate the process of *solidification*: How collective identities become activated and politically relevant.

[16] Iraq has a majority Shi'a population, but was ruled by Saddam Hussein (a Sunni) until 2003. With

groups which do not benefit from official protection that form the basis of my analysis.

4 THEORETICAL FRAMEWORK

4.1 MODEL 1: CONTAGION AND INSTITUTIONAL CONSTRAINTS

My first model is based upon Besley and Persson (2011). In the Besley-Persson (B-P) framework, society is divided into two large ethnic coalitions, the incumbency (I) and opposition (O). The opposition can challenge the state by building up an insurgency, L^O; while the state defends its power through the national army, L^I. These are the choice variables of the model. Each is assumed to be non-negative. Three key stochastic variables determine the state of the system: Resource rents (R), institutions (θ) and transnational spillovers (β) The state draws its revenue from resource rents, which follow a stochastic distribution on support $R \in [R_l, R_h]$. At the start of the period, redistribution of surplus rents between the two groups is exogenously determined by institutional constraints inherited from the previous period. The strength of constraints on redistribution is measured by the institutional parameter$\theta \in [0, \frac{1}{2}]$. A low value of θ signifies weak institutions that generate political and economic inequality, as the state can claim the greater part of resource rents and power for its own use; while $\theta = \frac{1}{2}$ represents strong institutional constraints and perfect equality between the two groups.

The stochastic parameter $\beta \in [0, 1]$ is this paper's key innovation to the B-P framework. The parameter measures the strength of spillovers from conflicts involving transnational ethnic kin of the opposition. Positive spillovers alter the preferences of group O over redistribution, causing it to demand a new and higher level of institutional inclusiveness given by $\theta' \in [\frac{1}{2}, 1]$.[17] They also augment the fighting effort of the opposition. Both of these

the aid of Saudi Arabia and other powerful allies, the Sunni Bahraini monarchy has retained control over a Shi'a majority. Syria, with its majority Sunni population, is ruled by the Assad dynasty, which is Alawite (a form of Shi'ism). Religious minorities under the protection of Bashar al-Assad's regime, such as Shi'as, Christians and Druze, fear falling prey to Sunni violence should the regime fall, and tend to support it for this reason. See Blanford (2015) and Lynch (2014).

[17] The new level of θ can be interpreted as a latent, long-standing grievance that only becomes activated (politically relevant) in the presence of spillovers and the demonstration effect. The value of this new parameter is therefore treated as predetermined.

effects are mediated by β. In employing a single β, we are hypothesizing that spillovers exercise a common impact upon the size of the insurgent army and its redistribution preferences.

Three states of the world are possible: Peace ($L^I = L^O = 0$), repression ($L^I > 0, L^O = 0$) and civil war ($L^I > 0, L^O > 0$).[18] The probability that the opposition can successfully seize power is given by the contest success function $\gamma \in (0, 1)$. By assumption, the C.S.F. is increasing and concave in the fighting effort of the opposition, and decreasing and convex in that of the incumbent:

(1) $\gamma_O > 0$, $\gamma_{OO} < 0$, $\gamma_I < 0$, $\gamma_{II} > 0$

These assumptions impose that likelihood of a successful insurgency rise as the opposition increases the size of its forces, and fall as the incumbent's army grows. There are decreasing returns to fighting effort.

Each iteration of the model has six stages:

1. The value of resource rents R and of spillover parameter β is realized.
2. Group O demands level θ' of redistribution.
3. Groups I and O simultaneously choose their fighting efforts, L^I and L^O.
4. With probability γ, group O attains power; with probability $1 - \gamma$, group I retains it.
5. The new incumbent assumes power and decides the level of public spending (G), the

[18]Besley and Persson exclude a fourth state, rebellion without repression ($L^I = 0, L^O > 0$), through opportune constraints on the parameter space. In the proof of Prop. 2, I discuss how adjusting this constraint impacts the model. The exclusion of rebellion without repression from our model is justifiable on the basis that a large share of revolts begin as peaceful protests, only evolving into a violent insurgency once a regime begins to suppress dissent through bloodshed. Heydemann and Leenders (2014) argue that violent repression of Arab Spring movements emerged largely in the wake of the West's "betrayal" in March 2011 of Hosni Mubarak, president of Egypt and formerly a close ally to the Americans. Not only protest movements, but ruling regimes as well benefit from the demonstration effect, updating their beliefs after observing events abroad. The harsh treatment of Mubarak "eroded perceptions that credible exit strategies existed, offering credible guarantees of future security" (*Ibid*). Bashar al-Assad's refusal to cede power in Syria, and the civil war that has ensued, are at least partly traceable to his concern about suffering the same fate as his fellow autocrat. The killing of Muammar al-Gaddafi, leader of Libya, in October 2011 only cemented these fears. The ability of Omar Hassan al-Bashir to evade such an end and remain president of Sudan, despite indictment by the International Criminal Court in 2010 on counts of genocide and crime against humanity, instead provided a valuable lesson on the usefulness of defying Western demands (Simons, 2010).

extent of redistribution (private transfers T^I and T^O), and future political inclusiveness (θ).

6. Payoffs are realized, consumption takes place, and a new period begins.

4.1.1 UTILITY AND FIGHTING EFFORT

The Subgame Perfect Equilibrium of the dynamic game is solved for by backward induction. In the final stage, each group derives utility from two sources, the public good (G) and private consumption (c^J):

$V^J = H(G) + c^J \ J \in \{O, I\}$

Private consumption is provided by group specific transfers, T^I and T^O. In stage 5, the government must determine the extent of public spending and private transfers. It allocates the budget between G, T^J and L^I so as to satisfy the constraint:

$R - \sum_J T^J - G - L^I \geq 0$

The provision of public goods is determined according to the rule:

$G = min\left\{\hat{G}, R - L^I\right\}$

where $\hat{G} = H_G^{-1}(1-\theta)$ is obtained by plugging the budget constraint into the expression for V^J. The first term of the minimand is an interior solution, with positive private transfers; the second term is a boundary solution, when resource rents are insufficient to fund private transfers. When institutions are perfectly strong and inclusive ($\theta = \frac{1}{2}$), the public good is provided according to the rule $\widetilde{G} = H_G^{-1}(\frac{1}{2})$.

In the case of an interior solution, group O receives a share θ of available resources, while I receives $1-\theta$.

$T^I = (1-\theta)(R - G - L^I)$

$T^0 = \theta(R - G - L^I)$

However, this simple sharing rule is altered in the presence of positive spillovers from transnational ethnic conflict ($\beta > 0$): Group O develops new preferences over redistribution, given by the parameter $\theta' > \theta$. O's utility from the private transfer is discounted by a factor proportional to the distance between θ and θ', the actual and desired levels of redistribution.

$$T^O = \theta\left(1 - \beta d(\theta' - \theta)\right)\left(R - G - L^I\right)$$

Note the first key asymmetry between the two sides of the conflict: Only the utility of the opposition is impacted by spillovers, as the incumbent maintains the same preferences as before.

In stage 4, the outcome of the conflict is determined by the function γ. In stage 3, the government and opposition build their armies. The government chooses L^I to maximize expected utility:

$$\hat{V^I} = H(G) + \left[\left(1 - \gamma(L^O, L^I)\right)T^I + \gamma(L^O, L^I)T^O\right]\left(R - G - L^I\right)$$

While the opposition chooses L^O to maximize:

$$\hat{V^O} = H(G) + \left[\gamma(L^O, L^I)T^I + \left(1 - \gamma(L^O, L^I)\right)T^O\right]\left(R - G - L^I\right) - L^O\left(1 + \beta \sum_K g_{OK}L_K\right)$$

where $g_{OK} \in \{0, 1\}$are coefficients from an adjacency vector G_{Nx1}measuring whether or not group $K \in \{1, ..., N\}$ is an ethnic ally of the opposition. The fighting effort of each allied ethnic group is given by L_K. Note that the incumbent's army is paid for out of the public budget, while the opposition must fund fighting effort from its resources. This is the second key asymmetry of the model. From the expression for $\hat{V^O}$ and for $\beta > 0$, it is clear that the size of the insurgent force rises and falls with the fighting effort of ethnic allies abroad. But what about the size of the government's army? It also moves positively with fighting abroad, as we demonstrate below.

4.1.2 EQUILIBRIUM

A final step in solving for the equilibrium is to show that in the case of positive transnational spillovers, the fighting effort of minorities abroad, the opposition and the government are indeed strategic complements, mutually reinforcing each other. This step is necessary because we will characterize the equilibrium as three increasingly violent states of the world, ordered by β. The realization of the spillover parameter in stage (1) ultimately determines the fighting effort of each side of the conflict, and whether we end up in a state of peace, repression or civil war.

PROPOSITION 1 (*Strategic complements*) Let $(\hat{L}^I, \hat{L}^O)$ be a Subgame Perfect Equilibrium of the conflict game with positive transnational spillovers $(\beta > 0)$. Then there exist two thresholds such that

Assume the following two conditions both hold:

(a) $\frac{1}{2} > \theta - \frac{1}{2}\theta\beta d(\theta' - \theta) \Rightarrow \frac{1}{2} > \theta$

(b)γ_{IO}>γ_I

If, in addition to (a) and (b) we have:

(c) $(\gamma_I + \gamma_{IO})(1 - 2\theta + \theta\beta d(\theta' - \theta)) > -(1 + \beta \sum_K g_{OK} L_K)\left(\frac{-\gamma_{II}}{\gamma_O} - (\frac{\gamma_{IO}}{\gamma_O})^2\right) - \gamma_I$

Then $\hat{L}^O$is increasing in β;

If, in addition to (a) and (b) we have:

(d) $\frac{-\gamma_{OO}\gamma_{II}}{\gamma_O^2} + \frac{\gamma_{O(1-2\theta+\theta\beta d(\theta'-\theta))}}{1+\beta\sum_K g_{OK} L_K} > \frac{\gamma_{IO}}{\gamma_O}$

Then $\hat{L}^I$is increasing in β.

Moreover, $\hat{L}^O$and $\hat{L}^I$are strategic complements.

PROOF. See appendix.

Condition (a) places an upper limit on the institutional parameter that always holds (save in the case of perfect equality). Because we have assumed $\gamma_I < 0$, (b) holds whenever $\gamma_{IO} > 0$. This is equivalent to assuming that L^I and L^O are strategic complements. Thus for positive spillovers and under these four conditions, the fighting effort of ethnic allies abroad, of domestic minorities and of the state are all increasing together. We are ready to characterize the equilibrium in terms of threshold values of the key parameter, β.

PROPOSITION 2 (*Trigger points for violence*) Let $(\hat{L}^I, \hat{L}^O)$ be a Subgame Perfect Equilibrium of the conflict game with positive transnational spillovers $(\beta > 0)$. Then there exist two thresholds β^I and β^O:

$$\beta^I = \frac{-\gamma(0,0)(1-2\theta)+1-\theta+\frac{\gamma_I(0,0)}{\gamma_O(0,0)}}{\theta\gamma(0,0)d(\theta'-\theta)-\frac{\gamma_I(0,0)}{\gamma_O(0,0)}\sum_K g_{OK} L_K}$$

$$\beta^O > \beta^I$$

Such that:

1. For $\beta < \beta^I$ there is peace with $\hat{L}^I = \hat{L}^O = 0$
2. For $\beta \in \left[\beta^I, \beta^O\right]$ there is repression with $\hat{L}^I > 0, \hat{L}^O = 0$
3. For $\beta > \beta^O$ there is civil war with $\hat{L}^I > 0, \hat{L}^O > 0$

PROOF. See appendix.

Three outcomes occur in equilibrium. When the spillover parameter β falls below the lower bound, β^I, neither group has the incentive to fight. For the intermediate case ($\beta \in \left[\beta^I, \beta^O\right]$), only the state engages in violence. For $\beta > \beta^O$, the opposition rebels violently and civil war results. Studying the lower bound for violence, several facts become immediately apparent. First, the threshold is decreasing in the fighting effort of foreign ethnic minorities (represented by the term $\sum_K g_{OK} L_K$). This holds because γ_I<0 by assumption. Second, it is falling in the probability of a peaceful transition of power (given by $\gamma(0,0)$). Third, it falls as the distance between the opposition's preferred level of redistribution (θ') and actual θ rises, somewhat counterintuitively. This result suggests that the state's incentives to engage in violent repression decline as the grievances of the marginalized ethnic group grow.

The impact of the institutional framework upon the likelihood of violence appears quite ambiguous. When we turn our attention to the direct impact of θ upon the β^I, we find that the threshold rises in θunder two conditions: $\gamma(0,0) > 0$, so that a peaceful transition to power is very likely; and $\partial/\partial\theta\{d(\theta' - \theta\}$ small, so that the distance between preferred and effective redistribution falls only very slowly.[19] When a peaceful transition is improbable or the distance falls quickly, the threshold is increasing in the level of redistribution. Thus institutional constraints can actually increase the incentives of the state to engage in repression, when spillovers are present.

I now present the second model of transnational contagion.

[19]The numerator of $\partial\beta^I/\partial\theta$ is given by $(2\gamma(0,0) - 1)\left[\theta\gamma(0,0)d(\theta' - \theta) - \frac{\gamma_I}{\gamma_O}\sum g_{OK}L_K\right] + [\gamma(0,0)(1 - 2\theta) - 1 + \theta - \gamma_I]\left[d(\theta' - \theta) + \theta\frac{\partial}{\partial\theta}d(\theta' - \theta)\right]$

4.2 MODEL 2: CONTAGION AND *DE FACTO* POWER

My second model is based upon Morelli and Rohner (2015). In the Morelli-Rohner (M-R) model, ethnic groups and natural resource rents are heterogeneously distributed across the national territory. The opposition (O) is concentrated in region 1, while the incumbent (I) is concentrated in region 2, so that $N_O^1/N_I^1 > N_O^2/N_I^2$ where N_J^L is the population of group J in region L. Resource endowments are given by R_1 (Region 1) and R_2 (Region 2). An unstable situation can unfold when the opposition inhabits the resource-rich region ($R_1 >> R_2$), so that its *de facto* power on the basis of resource wealth far outpaces its *de jure* power as determined by the political system.

There are three ordered states in equilibrium, divided by threshold values of θ. This is the parameter determining how surplus resource rents or power are redistributed between the two groups. In contrast to the first model, θ no longer represents an institutional constraint but a policy choice on the part of the incumbent. In choosing a low value of θ, the incumbent creates an unequal and unjust system; a high value engenders justice and equality between the two groups. The opposition bases its own strategy on the minimum value of θ that will dissuade it from violent rebellion.

The uniqueness of the M-R framework lies in the fact that both O and I can trigger war: The incumbent, by setting an upper bound $\bar{\theta}$ to redistribution; the opposition, by imposing a lower bound that exceeds $\bar{\theta}$. Depending upon the realization of the stochastic parameters, peace, secessionist war or all-out centrist war over the entire territory of the state are the resulting outcomes. Each side sustains the same cost of war: $q_C^O = q_C^I = q_C$ and $q_S^O = q_S^I = q_S$. The probabilities of victory—determined by the contest success functions p_C, for a centrist war, and p_S (secessionist)—are different for the two sides. We assume that $p_S > \frac{1}{2} > p_C$: The opposition is favored in a secessionist war; the state, in a centrist one, reflecting how the balance of power shifts as one moves across the national territory.

The stochastic parameter $\beta \in [0, 1]$ is once again this paper's key innovation to the original theoretical framework. The parameter mediates the impact of spillovers from transnational ethnic warfare upon domestic security. When a shock of this nature is realized, the preferences of the opposition over redistribution are modified, so that the opposition is assigned a share θ but receives only $\theta(1-\beta d(\theta'-\theta))$ in utility. Moreover, the contest success functions are influenced in a positive way, an innovation I will discuss at the end of this section. This change mirrors the increase in the opposition's fighting effort that occurs in the first model as a consequence of spillovers. Spillovers thus introduce a second source of instability into the system, in addition to the heterogeneous distribution of resources across national territory that is the focus of the original M-R model. Through any of the several channels of contagion shown in Figure 1, spillovers increase the *de facto* power of the opposition even as its *de jure* power remains frustratingly constrained by the ruling ethnic coalition.[20]

Each iteration of the model has five stages:

1. The value of the spillover parameter β is realized.

2. Group I chooses the level of redistribution, θ.

3. The two groups choose simultaneously and non-cooperatively between peace, centrist war or secessionist war. Both groups must choose peace in order for this outcome to be realized.

4. If at least one group chooses centrist war, then the opposition wins with probability p_C. If secessionist war is chosen, the opposition wins with probability p_S. If the two groups choose different wars, the opposition emerges victorious with probability $p = \frac{1}{2}(p_C + p_S)$.

5. Payoffs are realized, consumption occurs, and a new generation is born.

4.2.1 UTILITY AND REDISTRIBUTION

The equilibrium concept is Subgame Perfect Equilibrium. We solve for it through backward induction, beginning with the last stage of the game. In the case of peace, national resource rents $(R_1 + R_2)$ are redistributed according to sharing rule θ:

$$\pi_p^O = \theta(1 - \beta d(\theta' - \theta))(R_1 + R_2)$$

$$\pi_P^I = (1 - \theta)(R_1 + R_2)$$

In centrist war, the victor appropriates all resource wealth, leaving the loser with nothing.

[20] Though the contest function is not explicitly modeled in the first framework, nor is fighting effort in this framework, there is likely to be a powerful positive correlation between fighting effort and probability of victory in every war (given a broad enough definition of fighting effort). We are thus exploring essentially the same effect in both frameworks.

$\pi_C^O = p_C(R_1 + R_2) - q_C$

$\pi_C^I = (1 - p_C)(R_1 + R_2) - q_C$

If a secessionist war is successful, each ethnic group seizes control of rents from its own territory. Otherwise the opposition receives nothing.

$\pi_S^O = p_S R_1 - q_S$

$\pi_S^I = (1 - p_S)(R_1 + R_2) + p_S R_2 - q_S$

In stage 3, each group compares expected payoff under the different scenarios in order to formulate threshold values for θ. Comparing payoffs under peace versus centrist war (its preferred conflict), we find that the maximum amount the state is willing to pay in order to avoid conflict is given by:

$\bar{\theta} = p_C + \frac{q_C}{R_1+R_2}$

The least amount that will dissuade the opposition from centrist war is given by:

$\underline{\theta}_C = \frac{1}{1-\beta d(\theta'-\theta)} \left(p_C - \frac{q_C}{R_1+R_2}\right)$

The opposition is dissuaded from secessionist war by:

$\underline{\theta}_S = \frac{1}{1-\beta d(\theta'-\theta)} \left(p_S \frac{R_1}{R_1+R_2} - \frac{q_S}{R_1+R_2}\right)$

Comparing the thresholds in the case of spillovers ($\beta > 0$) vs. no spillovers ($\beta = 0$), it is clear that transnational contagion instigates a far higher frequency of conflict. The government's upper bound on redistribution is unaltered, as its incentives remain unchanged; but the opposition's two lower bounds are both increasing in θ. The parameter space in which bargaining between the two groups can be successfully concluded is thus restricted.

Before moving to the characterization of the equilibrium, we can determine conditions under which different states of the world occur by comparing different threshold values. I analyze when centrist war can be averted in favor of peace (a), when secessionist can (b), when the opposition prefers centrist over secessionist conflict (c), and when the state prefers a centrist war (d). These different conditions on the parameters determine the outcomes that can occur in equilibrium.

a) $\bar{\theta} > \underline{\theta}_C$ under spillovers when $2 > \beta d(\theta' - \theta)$: To avoid conflict, it is necessary to impose upper limits on β and on the distance between θ' and θ. Without spillovers, the condition always holds; the opposition never favors centrist war over peace.

b) $\bar{\theta} > \underline{\theta}_S$ under spillovers when

$p_C(R_1 + R_2) - \frac{p_S R_1}{1-\beta d(\theta'-\theta)} > \frac{-q_S}{1-\beta d(\theta'-\theta)} - q_C$

This always holds for $1 > \beta d(\theta' - \theta)$. Without spillovers, it holds for:

$p_C(R_1 + R_2) - p_S R_1 > -q_S - q_C$

This is always true when $p_C(R_1 + R_2) \geq p_S R_1$.

c) $\underline{\theta}_C > \underline{\theta}_S$ when

$p_C(R_1 + R_2) - p_S R_1 > q_C - q_S$

That is, when the LHS is large and positive $(p_C(R_1 + R_2) >> p_S R_1)$ or the RHS is large and negative $(q_S >> q_C)$. Note that spillovers do not impact preferences between different wars, but only the desire for war or peace; this is also true for the government.

d) $\pi_C^I > \pi_S^I$ when $p_S R_1 >> p_C(R_1 + R_2)$ or $q_S >> q_C$.

Comparing the last two results, it is clear that the state and opposition prefer different conflicts based upon the value of $p_S R_1$, the expected gain of the opposition in the case of a secessionist war. Based upon cost, they prefer the same type of conflict; this occurs due to our initial assumptions on q_C and q_S.

In stage 2, the state determines the share of resources to redistribute to the opposition, given the thresholds carried over from stage 3 and the outcome it wishes to induce.

4.2.2 EQUILIBRIUM

We are now ready to compute when each of the three states occur in equilibrium. Two results from the analysis in the previous section— parts a) and b), stating when centrist and secessionist wars can be averted— lead to the main results of our characterization of equilibrium.

PROPOSITION 3 (*Peace and War*) Let $\left(\bar{\theta}, \underline{\theta}_C, \underline{\theta}_S\right)$ be a Subgame Perfect Equilibrium of the conflict game with positive transnational spillovers $(\beta > 0)$.

For $2 > \beta d(\theta' - \theta)$, $\bar{\theta} > \underline{\theta}_C$ and centrist war can always be averted.

For $1 > \beta d(\theta' - \theta)$, $\bar{\theta} > \underline{\theta}_S$ and secessionist war can always be averted.

In other cases, war can be avoided if the cost of war, represented by q_C and q_S, is high enough.

PROOF. See appendix.

War can be avoided in three cases. First, the spillover parameter β must be small. Second, the distance between the preferred and effective levels of redistribution cannot be too large, as both $\underline{\theta}_C$ and $\underline{\theta}_S$ are increasing in this term. This finding is in direct contrast to the equilibrium results from the first model. Finally, war must be a costly action. The opposition's lower bounds on redistribution rise in q_C and q_S, while the upper bound set by the state falls. This last finding is curious, for we have assumed a homogenous cost function for both sides of the war. One can envision scenarios under which either side undertakes a devastatingly costly action in the hope that the cost will destroy its enemy first (as in wars of attrition such as the Eastern Front of World War II). Such wars of attrition are excluded from this framework. Each group shapes its own behavior based upon the expected cost of war, without accounting for the fact that the other group will have to pay the exact same cost. Attrition settings could be built into the model by introducing complicated cost functions that incorporate expectations about the relative resistance of the other side.

In a large range of cases, war cannot be avoided. Centrist war becomes a certainty for $\beta d(\theta' - \theta) \geq 2$; secessionist war is possible when $\beta d(\theta' - \theta) \geq 1$. Secessionist war is therefore more difficult to avoid. This makes sense intuitively, because secessionist wars are generally preferred by the opposition, and spillovers have a larger impact upon ethnic groups in this coalition. The range of parameter values under which war occurs in the presence of spillovers is far wider than the range without them, as the proof to Prop. 3 demonstrates.

From Prop. 3, it is clear that the impact of spillovers on the risk of war is large and significant. This impact is entirely due to the shift in preferences of the opposition over redistribution. I hypothesize that the fighting effort of transnational ethnic kin exercises a second fundamental impact upon the domestic system. Because the *de facto* power of the opposition is enhanced by spillovers, it is likely that spillovers will positively influence the likelihood of a successful insurgency. This effect is parallel to the change induced in fighting effort in the first model. I follow M-R in adopting a ratio formulation for the contest function:

$$p_C = \frac{N_O y_O}{N_O y_O + N_I y_I + \lambda - \beta \sum_K g_{OK} L_K}$$

$$p_S = \frac{N_O^1 y_O}{N_O^1 y_O + N_I^1 y_I + \lambda/(1-\phi) - \beta \sum_K g_{OK} L_K}$$

Where y_J is non-resource income of group J; λ represents the added advantage of the incumbent from its monopoly over state power; and ϕ is an increasing function of the distance between the state capital and the opposition's homelands. In subtracting spillovers from the λ parameter introduced by M-R, we express how state power can be subverted by foreign influences. Indeed, this difference could become negative for particularly influential transnational ethnic kin. With this modification to the contest function the threshold for war falls still lower, making violence all but inevitable in many cases.

5 DISCUSSION AND CONCLUSIONS

In this paper, I have developed a framework for modelling the spillover of conflict across national borders through transnational ethnic ties. I have shown that when spillovers are present, a country finds itself at far greater risk of bargaining failure and conflict. I argue that this failure in the ability of the incumbent and opposition to reach an agreement arises from a change in a political system's balance of power. When ethnic groups allied with the opposition rebel against their governments, the opposition becomes stronger and more determined to challenge the domestic status quo. The ruling regime feels threatened and responds accordingly: With concessions, if it possesses adequate resources; with violence if it does not.

The logical next step in my analysis would be to specify the reduced form of the model and identify the parameters to be estimated. Ideally, this framework would be flexible enough to capture the determinants of war in any region of the world impacted by secessionist uprisings and civil wars, from Southeast Asia to South America and beyond. The Kurdish communities spanning the Middle East, for example, would provide an ideal testing ground for my hypotheses regarding the power of transnational kin to transform the arena of domestic politics. Brathwaite (2014) has performed a qualitative analysis along these lines.

My framework challenges the tradition of means-opportunity thinking in economic models of civil conflict. Resource scarcity and resource wealth play a crucial role in both the models I have set forth, but I believe that in order to understand the complicated processes of peace and strife occurring within a country, it is necessary to look beyond state borders and conceive of the problem in new ways. The interconnectedness of seemingly disparate domestic phenomena is something that sociology and political science have long understood. The phenomenal worldwide expansion of Islamic State, as of al-Qaeda before it, suggests the power of identity-based thinking. By introducing ethnicity and collective identities into economic models of war, we have an exciting opportunity to learn more about a fundamental force that drives communities together and confers strength far beyond what the opportunity structure would suggest.

6 APPENDIX

6.1 PROOF OF PROPOSITION 1

Denoting available resources (R-G) as Z, the F.O.C. for L^I and L^O are given by:

$$-(1-\theta)(1-\gamma)-\theta\gamma(1-\beta d(\theta'-\theta))-\gamma_I(1-2\theta+\theta\beta d(\theta'-\theta))(Z-L^I)=0$$

$$\left[\gamma_O(1-2\theta+\theta\beta d(\theta'-\theta))\right](Z-L^I)-1-\beta\sum_K g_{OK}L_K=0$$

Where the dependence of γ upon L^I and L^O has been suppressed in order to simplify notation. We want to show that L^I and L^O are both increasing in β. We can do this by taking the total differential of the system and rewriting it in matrix form:

$$\begin{matrix} \frac{-\gamma_{II}}{\gamma_O}(1+\beta\sum_K g_{OK}L_K)+\gamma_I(2-2\theta+\theta\beta d(\theta'-\theta)) & \frac{\gamma_{IO}}{\gamma_O}(1+\beta\sum_K g_{OK}L_K)-\gamma_O(1-2\theta+\theta\beta d(\theta'-\theta)) \\ \frac{\gamma_{IO}}{\gamma_O}(1+\beta\sum_K g_{OK}L_K)-\gamma_O(1-2\theta+\theta\beta d(\theta'-\theta)) & \frac{\gamma_{OO}}{\gamma_O}(1+\beta\sum_K g_{OK}L_K) \end{matrix}$$

Denote the determinant of the coefficient matrix by Ω:

$$\left[\frac{-\gamma_{OO}\gamma_{II}}{\gamma_O^2}(1+\beta\sum_K g_{OK}L_K)^2+\frac{\gamma_{OO}\gamma_I}{\gamma_O}(1+\beta\sum_K g_{OK}L_K)(2-2\theta+\theta\beta d(\theta'-\theta))\right]+\left[\frac{\gamma_{IO}}{\gamma_O}(1+\beta\sum_K g_{OK}L_K\right.$$

$$0$$

Applying Cramer's rule and simplifying, we obtain the derivatives of L^I and L^O with respect to β:

$$\frac{\partial L^I}{\partial\beta}=\frac{\sum_K g_{OK}L_K\left(1+\beta\sum_K g_{OK}L_K\right)\left[\frac{-\gamma_{OO}\gamma_{II}}{\gamma_O^2}-\frac{\gamma_{IO}}{\gamma_O}+\frac{\gamma_O(1-2\theta+\theta\beta d(\theta'-\theta))}{1+\beta\sum_K g_{OK}L_K}\right]}{\Omega}$$

$$\frac{\partial L^O}{\partial\beta}=\frac{\sum_K g_{OK}L_K\left[\left(1+\beta\sum_K g_{OK}L_K\right)\left(\frac{-\gamma_{II}}{\gamma_O}-(\frac{\gamma_{IO}}{\gamma_O})^2\right)+(\gamma_I+\gamma_{IO})(1-2\theta+\theta\beta d(\theta'-\theta))+\gamma_I\right]}{\Omega}$$

We will first consider the fighting effort of the opposition. When is L^O increasing in β? Recalling our assumptions (1) on the contest function ($\gamma_O>0,\gamma_{OO}<0,\gamma_I<0,\gamma_{II}>0$), the first and third added terms inside the square brackets are negative. Therefore, the middle term must be positive. This occurs when the two multiplied factors are either both negative or both positive.

Case 1) Two positive factors; both (a) and (b) must hold:

(a) $\frac{1}{2}>\theta-\frac{1}{2}\theta\beta d(\theta'-\theta)\Rightarrow\frac{1}{2}>\theta$

which always holds (save for the case of perfect equality) because $\theta \in [0, \frac{1}{2}]$.

(b) $\gamma_{IO} > \gamma_I$

This holds whenever $\gamma_{IO} > 0$ (strategic complements)

Case 2) Two negative factors: both (c) and (d) must hold:

(c) $\frac{1}{\theta} < 2 - \beta d(\theta' - \theta) \Rightarrow \frac{1}{2} < \theta$

(d) $\gamma_{IO} < \gamma_I$

However, (c) never holds. Therefore both factors must be positive. This is still not enough to ensure that fighting effort is increasing in β, however. The second term must be larger in magnitude than the sum of the first and third:

(e) $(\gamma_I + \gamma_{IO})\,(1 - 2\theta + \theta\beta d(\theta' - \theta)) > -\,(1 + \beta \sum_K g_{OK} L_K) \left(\frac{-\gamma_{II}}{\gamma_O} - (\frac{\gamma_{IO}}{\gamma_O})^2\right) - \gamma_I$

Conditions (a), (b) and (e) are jointly necessary and sufficient for $\partial L^O / \partial \beta > 0$.

Turning to the fighting effort of the government, $\partial L^I / \partial \beta > 0$ when the term in square brackets is greater than zero. Under the assumptions on γ, the first term inside the square brackets is always positive, as is the denominator of the third term. We have also assumed $\gamma_O > 0$. Thus the third term is positive whenever condition (a) holds. Given (b), the second term is negative. Positivity of the derivative is assured when:

(f) $\frac{-\gamma_{OO}\gamma_{II}}{\gamma_O^2} + \frac{\gamma_O(1 - 2\theta + \theta\beta d(\theta' - \theta))}{1 + \beta \sum_K g_{OK} L_K} > \frac{\gamma_{IO}}{\gamma_O}$

To summarize, if (a), (b) and (e) hold then $\partial L^O / \partial \beta > 0$ (strategic complementarities in foreign minority and domestic opposition effort). If (a), (b) and (f) hold then $\partial L^I / \partial \beta > 0$ (strategic complementarities in foreign minority and domestic government effort). Because we have assumed $\gamma_{IO} > 0$, $L^I$and L^Oincrease together.

6.2 PROOF OF PROPOSITION 2

Four scenarios are possible: 1) peace 2) only violent rebellion; 3) only violent repression and 4) rebellion and repression (civil war). B-P exclude the case of unchallenged rebellion by restricting the parameter space such that the F.O.C. with respect to L^I is greater than the F.O.C. with respect to L^O, when $L^I = L^O = 0$. The state thus has a larger marginal utility from breaking the peace by engaging in repression. If the two F.O.C. are equal for $L^I = L^O = 0$, then there is a single threshold and society jumps from peace directly into civil war.

We therefore have three ordered states of the world, 1) -> 3) -> 4) which we move between according the value of the new key parameter BETA.

Define $L^I = \hat{L^I}(\beta)$ when $L^O = 0$. We find the lower threshold,β^I, using $\partial L^I/\partial\beta$ and setting $\hat{L^I}(\beta) = 0$.

This the threshold such that the government moves from peace to repression. From $\partial L^I/\partial\beta$:

$$-\left(1-\theta\right)\left(1-\gamma(0,0)\right)-\theta\gamma(0,0)\left(1-\beta d(\theta\prime-\theta)\right)-\frac{\gamma_I(0,0)}{\gamma}\left(1+\beta\sum_K g_{OK}L_K\right)=0$$

where we have substituted for $Z - L^I$ from $\partial L^O/\partial\beta$. Solving for β,we obtain

$$\beta^I = \frac{-\gamma(0,0)(1-2\theta)+1-\theta+\frac{\gamma_I(0,0)}{\gamma_O(0,0)}}{\theta\gamma(0,0)d(\theta'-\theta)-\frac{\gamma_I(0,0)}{\gamma_O(0,0)}\sum_K g_{OK}L_K}$$

We can define β^O,the threshold such that opposition has positive fighting effort, implicitly from (2):

$$\gamma_O(0,\hat{L^I}(\beta))\left(1-2\theta+\theta\beta d(\theta'-\theta)\right)\left(Z-L^I\right)-1-\beta\sum_K g_{OK}=0$$

β^O is threshold for the realization of civil war. I have not yet shown that $\beta^O > \beta^I$. This is a necessary step for the future development of the model.

If we were instead to exclude 3) repression in the absence of rebellion, then we would move from 1)->2)->4) based upon the value of β.

To exclude 3), we require $\partial L^O/\partial\beta > \partial L^I/\partial\beta$ when $L^I = L^O = 0$

Define $L^O = \tilde{L^O}(\beta)$ when $L^I = 0$. We find the lower threshold $\beta^{O'}$ using F.O.C. (2) and setting $\tilde{L^O}(\beta) = 0$

This is the threshold such that the opposition moves from peace (or peaceful protests) to violent rebellion. From (2):

$$\gamma_O(0,0)\left(1-2\theta+\theta\beta d(\theta'-\theta)\right)Z-1-\beta\sum_K g_{OK}L_K=0$$

$$\beta^{O'} = \frac{\gamma(0,0)(1-2\theta)Z-1}{\sum_K g_{OK}L_K-\gamma_O(0,0)\theta d(\theta'-\theta)Z}$$

We define $\beta^{I'}$, the threshold such that the government has positive effort, implicitly from (1)

$$-(1-\theta)\left(1-\gamma(\tilde{L^O}(\beta),0)\right)-\theta\gamma(\tilde{L^O}(\beta),0)\left(1-\beta d(\theta'-\theta)\right)-\frac{\gamma_I(\tilde{L^O}(\beta),0)}{\gamma_O(\tilde{L^O}(\beta),0)}\left(1+\beta\sum_K g_{OK}L_K\right)$$

$\beta^{I'}$ is the threshold for civil war in this case. It remains to be shown that $\beta^{I'} > \beta^{O'}$.

6.3 PROOF OF PROPOSITION 3

There are four cases for us to consider:

1) $\bar{\theta} > \underline{\theta}_C > \underline{\theta}_S > 0$

2a) $\bar{\theta} > \underline{\theta}_S > \underline{\theta}_C > 0$ and $\pi_S^I > \pi_C^I$

2b) $\bar{\theta} > \underline{\theta}_S > \underline{\theta}_C > 0$ and $\pi_C^I > \pi_S^I$

3) $\underline{\theta}_S > \bar{\theta} > \underline{\theta}_C > 0$

4) $\underline{\theta}_C > \bar{\theta} > \underline{\theta}_S > 0$

Case 4) never occurs in the absence of spillovers. In case 3), no surplus-sharing is possible such that conflict can be avoided, as Morelli and Rohner demonstrate for the no-spillover case. Bargaining failure in the no-spillover case also makes conflict inevitable for positive spillovers. This occurs because fighting effort is non-decreasing in transnational spillovers under the assumption of strategic complementarities. Therefore, whenever $\bar{\theta} < \underline{\theta}_C$ or $\bar{\theta} > \underline{\theta}_S$ war cannot be avoided.

We now turn our attention to cases 1), 2a) and 2b), when bargaining between the two groups over redistribution (concessions from the state to the opposition) might be possible.

Case 1) $\bar{\theta} > \underline{\theta}_C > \underline{\theta}_S > 0$

Expected profits without concessions:

$$E[\pi^I(\theta = 0)] = (1-p_C)(R_1+R_2) - q_C$$

Expected profits with concessions:

$$E[\pi^I(\theta = \underline{\theta}_C)] = (1-\underline{\theta}_C)(R_1+R_2)$$

The state cedes concessions when:

$q_C > (R_1 + R_2)\left(1 - \beta d(\theta' - \theta) - p_C + 2\beta d(\theta' - \theta)p_C\right)$

That is, when the cost of centrist warfare exceeds a certain threshold.

In the absence of spillovers:

$q_C > (R_1 + R_2)(1 - p_C)$

2a) $\bar{\theta} > \underline{\theta}_S > \underline{\theta}_C > 0$ AND $\pi_S^I > \pi_C^I$

Expected profits without concessions:

$E[\pi^I(\theta = 0)] = (1 - p_S)(R_1 + R_2) - q_S$

Expected profits with concessions:

$E[\pi^I(\theta = \underline{\theta}_S)] = (1 - \underline{\theta}_S)(R_1 + R_2)$

The state cedes concessions when:

$q_S > \left(1 - \beta d(\theta' - \theta)\right)(R_1 - 2p_S R_1 + R_2) + p_S R_1$

That is, when the cost of secessionist warfare exceeds a certain threshold.

In the absence of spillovers:

$q_S > (1 - p_S)R_1 + R_2$

2b) $\bar{\theta} > \underline{\theta}_S > \underline{\theta}_C > 0$ AND $\pi_C^I > \pi_S^I$

Expected profits without concessions:

$E[\pi^I(\theta = 0)] = (1 - p)R_1 + R_2 - q_S + (1 - p)(R_1 + R_2) - q_C$

where $p = \frac{1}{2}p_S + \frac{1}{2}p_c$

Expected profits with concessions:

$E[\pi^I(\theta = \underline{\theta}_S)] = (1 - \underline{\theta}_S)(R_1 + R_2)$

The state cedes concessions when:

$\frac{q_S}{1-\beta d(\theta'-\theta)} + q_C > \frac{\beta d(\theta'-\theta)}{1-\beta d(\theta'-\theta)} pR_1 + (2 - 2p)R_1 + (2 - p)R_2$

The costs of both types of war together must exceed the threshold given by the RHS.

In the absence of spillovers:

$$q_S + q_C > (2 - 2p)R_1 + (2 - p)R_2$$

Conditions for successful bargaining are far more arduous in the case of spillovers, as the RHS of each inequality is augmented by one or more positive addends.

References

[1] Acemoglu, D. and J. A. Robinson. 2009. *Economic Origins of Dictatorship and Democracy*. Cambridge: Cambridge University Press.

[2] Aragon, F.M., P. Chuhan-Pole and B.C. Land. 2015. "The Local Economic Impacts of Resource Abundance: What Have We Learned?" World Bank Policy Research Working Paper No. 7263.

[3] Belhaj, M., Y. Bramoullé and F. Deroïan. 2014. "Network Games Under Strategic Complementarities." *Games and Economic Behavior*, 88: 310–319.

[4] Besley, T. and T. Persson. 2011. "The Logic of Political Violence," *Quarterly Journal of Economics*, 126(3): 1411-1445.

[5] *Black's Law Dictionary: Free Online Legal Dictionary*. 1910. 2nd ed, available at: http://thelawdictionary.org/.

[6] Blanford, N. 2015. "For Syria's Druze, survival hinges on choosing the right ally." *Al Jazeera America*, July 2, available at: http://america.aljazeera.com/articles/2015/7/2/druze-struggle-for-survival-in-syria.html.

[7] Braithwaite, A. 2010. "Resisting Infection: How State Capacity Conditions Conflict Contagion." *Journal of Peace Research*, 47(3): 311-319.

[8] Bramoullé, Y., R. Kranton and M. D'Amours. 2014. "Strategic Interaction and Networks." *American Economic Review*, 104(3): 898-930.

[9] Brathwaite, K.J.H. 2014. "Repression and the Spread of Ethnic Conflict in Kurdistan." *Studies in Conflict & Terrorism*, 37: 473-491.

[10] Brownlee, J., T. Masoud and A. Reynolds. 2015. *The Arab Spring: Pathways of Repression and Reform*, 1st ed., Oxford: Oxford University Press.

[11] Buhaug, H. 2010. "Climate Not to Blame for African Civil Wars." Proc. Natl. Acad. Sci. USA, 107(38).

[12] Buhaug, H. and K.S. Gleditsch. 2008. "Contagion or Confusion? Why Conflicts Cluster in Space." *International Studies Quarterly,* 52(2): 215-233.

[13] Buhaug, H., N.P. Gleditsch and O.M. Theisen. 2008. "Implications of Climate Change for Armed Conflict." Paper prepared for the Social Dimensions of Climate Change Program. Washington, DC: World Bank, Social Development Department, available at: http://siteresources. worldbank. org/INTRANETSOCIALDEVELOPMENT/ Resources/SDCCWorkingPaper_Conflict.pdf.

[14] Burke, M.B., E. Miguel, S. Satyanath, J. A. Dykema and D. B. Lobell. 2009. "Warming increases the risk of civil war in Africa." Proc. Natl. Acad. Sci. USA, 106(49).

[15] Caselli, F., M. Morelli and D. Rohner. 2015. "The Geography of Interstate Resource Wars." *The Quarterly Journal of Economics,* 130: 267-315.

[16] Cederman, L.-E., L. Girardin, and K.S. Gleditsch. 2009. "Ethnonationalist Triads: Assessing the Influence of Kin Groups on Civil Wars." *World Politics*, 61: 403-437.

[17] Cederman, L.-E., A. Wimmer and B. Min. 2010. "Why Do Ethnic Groups Rebel? New Data and Analysis." *World Politics*, 62(1): 87-119.

[18] Cederman, L.-E., K.S. Gleditsch, I. Salehyana and J. Wucherpfennig. 2013. "Transborder Ethnic Kin and Civil War." International Organization, 67(2): 389-410.

[19] Collier, Paul and Anke Hoeffler. 2002. "Greed and Grievance in Civil War." World Bank, CSAE Working Paper 2002-01.

[20] Collier, P. and A. Hoeffler. 1998. "On the Economic Causes of War." Oxford Economic Papers, 50: 563-573.

[21] Dal Bó, E. and P. Dal Bó. 2011. "Workers, Warriors, And Criminals: Social Conflict In General Equilibrium." *Journal of the European Economic Association*, 9(4): 646-677.

[22] Danneman, N. and E.H. Ritter. 2014. "Contagious Rebellion and Preemptive Repression." *Journal of Conflict Resolution*, 58(2): 254-279.

[23] Diamond, J.M. 2006. "Collapse: How Societies Choose to Fail or Succeed." London: Penguin Books.

[24] Diamond, J.M. 1997. "Guns, Germs, and Steel: The Fates of Human Societies." New York: W. W. Norton & Company.

[25] Dube, O. and J. Vargas. 2013. "Commodity Price Shocks and Civil Conflict: Evidence From Colombia." *Review of Economic Studies*, 82(3).

[26] Esteban, J., M. Morelli and D. Rohner. 2015. "Strategic Mass Killings," forthcoming in *Journal of Political Economy*.

[27] Fearon, J.D. and D.D. Laitin. 2003. "Ethnicity, Insurgency and Civil War." *American Political Science Review*, 97: 75-90.

[28] Fosberg, E. "Polarizatin and Ethnic Conflict in a Widened Strategic Setting." *Journal of Peace Research*, 45: 283-300.

[29] Franke, J. and T. Ozturk. 2015. "Conflict Networks." *Journal of Public Economics*, 126(C): 104-113.

[30] Gilbert, Martin. 1985. T*he Holocaust: A History of the Jews of Europe During the Second World War*. New York: Henry Holt and Co.

[31] Gleditsch, N.P. 2007. "Transnational Dimensions of Civil War." Journal of Peace Research, 44(3): 293-309.

[32] Gleditsch, N.P. 2012. "Whither the Weather? Climate Change and Conflict." *Journal of Peace Research*, 49(1): 3-9.

[33] Gurr, T.R. 2000. *Peoples Versus States: Minorities at Risk in the New Century*. Washington, D.C.: United States Institute for Peace Press.

[34] Harari, M. and E. La Ferrara. 2013. "Conflict, Climate and Cells: A Disaggregated Analysis." CEPR Discussion Paper No. DP9277.

[35] Hegre, H., T. Ellingsen, S. Gates and N.P. Gleditsch. 2001. "Toward a Democratic Civil Peace? Democracy, Political Change and Civil War, 1816-1992." *American Poltical Science Review*, 95(1):33-48.

[36] Heydemann, S. and R. Leenders. 2014. "Authoritarian Learning and Counterrevolution." In *The Arab Uprisings Explained: New Contentious Politics in the Middle East*, M. Lynch, ed., New York: Columbia University Press.

[37] Hegre, H. and N. Sambanis. "Sensitivity Analysis of Empirical Results on Civil War Onset." *Journal of Conflict Resolution*, 50(4): 508-535.

[38] King, Jeremy. 2002. *Budweisers into Czechs and Germans*. Princeton, NJ: Princeton University Press.

[39] Konig, M.D., D. Rohner, M. Thoenig and F. Zilibotti. 2015. "Networks in Conflict: Theory and Evidence from the Great War of Africa." CEPR Discussion Paper No. DP10348.

[40] Lake, D.A. and D. Rothchild. 1998. *The International Spread of Ethnic Conflict: Fear, Diffusion, and Escalation*. Princeton, NJ: Princeton University Press.

[41] *The Long Way Home.* 1998. Documentary film, M.J. Harris dir.

[42] Lynch, M. 2014. *The Arab Uprisings Explained: New Contentious Politics in the Middle East*, New York: Columbia University Press.

[43] Mann, M. 2005. *The Dark Side of Democracy: Explaining Ethnic Cleansing*. Cambridge: Cambridge University Press.

[44] Maves, J. and A. Braithwaite. "Autocratic Institutions and Civil Conflict Contagion." *The Journal of Politics*, 75(2): 478-490.

[45] Meredith, M. 2007. *Mugabe: Power, Plunder, and the Struggle for Zimbabwe's Future*, New York: Public Affairs.

[46] Miguel, E., S. Satyanath and E. Sergenti. 2004. "Economic Shocks and Civil Conflict: An Instrumental Variables Approach." *Journal of Political Economy*, 112(4): 725-753.

[47] Morelli, M. and D. Rohner. 2015. "Resource Concentration and Civil Wars." *Journal of Development Economics*, 117: 32-47.

[48] Reuveny, Rafael. 2007. "Climate Change-Induced Migration and Violent Conflict." *Political Geography,* 26: 656-673.

[49] Robinson, J. A., D. Acemoglu, and S. Johnson. 2003. "An African Success Story: Botswana." In *In Search of Prosperity: Analytic Narratives on Economic Growth*, D. Rodrik, ed., 80-119. Princeton: Princeton University Press.

[50] Robinson, J. A, D. Acemoglu, and S. Johnson. 2005. "Institutions as a Fundamental Cause of Long-Run Growth." In *Handbook of Economic Growth* 1A: 386-472.

[51] Ross, M.L. 2013. "The Politics of the Resource Curse: A Review." Forthcoming in the Handbook on the Politics of Development, C. Lancaster and N. van de Walle, ed.

[52] Salehyan, I. and K.S. Gleditsch. 2006. "Refugee Flows and the Spread of Civil War." *International Organization,* 60(2): 335-366.

[53] Sambanis, N. 2002. "A Review of Recent Advances and Future Directions in the Quantitative Literature on Civil War." *Defense Economics*, 13(1): 215-243.

[54] Simons, Marlise. 2010. "International Court Adds Genocide to Charges Against Sudan Leader." *The New York Times*, July 12, available at: http://www.nytimes.com/2010/07/13/world/africa/13hague.html?_r=1.

[55] Stearns, J. 2012. *Dancing in the Glory of Monsters: The Collapse of the Congo and the Great War of Africa,* New York: Public Affairs.

[56] Tainter, J. 1988. *The Collapse of Complex Societies*. Cambridge: Cambridge University Press.

[57] van der Ploeg, F. 2010. "Natural Resources: Curse of Blessing?" CESifo Working Paper Resources and Environment No. 3125.

[58] VICE News. 2014. *The Islamic State*, documentary film, available at: https://news.vice.com/video/the-islamic-state-full-length.

[59] Wood, G. 2015. "What ISIS Really Wants." *The Atlantic*, March, available at: http://www.theatlantic.com/magazine/archive/2015/03/what-isis-really-wants/384980/.