War on Polio Eradication: Reaching the Hard-to-Reach

Laura Muñoz Blanco *

Trinity College Dublin

9th June, 2023

Abstract

Disease eradication is the mother of all global health efforts. And, we are closer than ever to ending polio, but the increasing number of internally displaced people (IDP) could pose new challenges. Which are the impacts of IDP inflows on host communities' polio incidence? To tackle this question, I use the mass displacement of the population from the conflict-affected Federally Administered Tribal Areas (F.A.T.A.) to other districts in Pakistan from 2008 to 2022. In a difference-in-differences approach, I compare the polio cases between host and non-host districts before and after 2007. I exploit the spatial distribution of districts with respect to the pre-colonial region of *Pashtunistan*'s border to define the host and non-host districts. I find that districts that received the IDP population increased the number of additional polio cases per 100,000 inhabitants by 40% over the mean incidence compared to non-host districts. There are three underlying mechanisms: overpopulated communities with low immunization rates, precarious health conditions, and the congestion of health services in host communities.

Keywords: internal displacement, infectious diseases, vaccines, Pakistan

JEL Classification: D60, I15, O15.

*PhD Candidate at Trinity College Dublin, Room 3014, Arts Building, Dublin 2, Ireland. I thank Martin Suesse for their guidance and time generosity. I thank María Alexandra Castellanos-Vasconez, Naomi Crowther, Lukas Delgado, Mélanie Gittard, Andrea Guariso, Nicola Mastrorocco, Danish Us-Salam, and Mirko Vintar for their helpful comments and suggestions. My special gratitude to the Global Polio Eradication Initiative from the World Health Organization for the data provided. All errors and omissions remain entirely my responsibility.

1 Introduction

Disease eradication is the mother of all global health efforts, as everyone worldwide can enjoy the well-being benefits of eradication. Smallpox is the only human disease to be eradicated worldwide. And we are closer than ever to ending polio, but outbreaks and challenges persist (UNICEF 2019).

The increasing number of internally displaced people (IDP) poses new challenges to eradicating polio worldwide. New conflicts are emerging in polio-affected countries, and with it, the number of people forced to flee. The forcibly displaced population climbed to 89.3 million by 2021. This figure is more than double the 42.7 million people who remained forcibly displaced in 2012. Internal displacement in 2021 was markedly higher than in recent years, corresponding to 66% of the total. This is the largest displaced population, but it is also one of the most vulnerable. The vast majority do not receive the medical care they need because they live in regions where the healthcare system has collapsed, or national governments do not protect their health rights. Therefore children in these fragile areas are at higher risk of contracting and spreading polio. (UNHCR 2022).

From a policy perspective, improving our knowledge of the role of IDPs in the incidence of polio is crucial for a better design and efficacy of preventive public policies. A polio-free world will save the global economy USD 45 billion in health costs within the next 20 years (UNICEF 2019). Additionally, eradicating polio could be the prelude to other diseases' eradication.

In this paper, I study the impacts of IDP inflows on polio incidence in host communities and evaluate some of the main potential channels. To tackle this question, I use the mass displacement of the 57% of the population from the conflict-affected Federally Administered Tribal Areas (F.A.T.A.) to other districts in Pakistan from 2008 to 2022. In a difference-in-differences approach, I compare the polio cases between host and non-host districts before and after 2007. Due to cultural and linguistic barriers, most IDPs settled within the historical region of *Pashtunistan*. I exploit the spatial distribution of districts with respect to the pre-colonial region of *Pashtunistan*'s border to define the host and non-host districts.

I find that districts that received the IDP population have experienced an increase in the probability of polio incidence compared to non-host districts. An IDP inflow increased the likelihood of at least one polio case by 5.3 percentage points (pp) and 0.007 additional cases per 100,000 inhabitants in the host districts. The estimates are statistically significant at the five per cent level. Although negligible, my findings represent an increase of 40% over the mean incidence.

I also look at the intensity of the inflow. To do so, I rely on districts closer to the F.A.T.A. border receiving more IDPs when the total yearly inflows of IDPs increase. More formally, as data on migration flows are unavailable at the district level, I construct a yearly district measure of predicted IDP inflows based on the interaction of the inverse distance to the F.A.T.A. border and the total yearly migration flows from F.A.T.A. to other regions in Pakistan. I show that an increase of one standard deviation in predicted inflows results in 0.001 additional polio cases per 100,000 inhabitants, corresponding to 20% of the mean incidence.

Why are the main effects meaningful? Polio, which only infects humans, has been eliminated in 193 countries. With the transmission of wild-type polio limited to Afghanistan and Pakistan, an official eradication declaration is in sight. In 2005, 28 cases were reported in Pakistan, compared to the 1,147 cases in 1997. Moreover, most host districts had zero or close to zero polio cases before 2007. Ultimately, the estimates I present in this paper capture the impacts of IDP inflows on keeping host communities away from eradicating polio rather than the effects on an extensive increase in polio cases.

There are three channels through which IDP inflow could potentially increase polio incidence in host communities. First, low vaccination rates at the national level increase the susceptibility to transmitting polio in overpopulated communities.

I propose three potential mechanisms by which IDP inflow could slow down polio eradication in host communities: a sudden increase in the population in communities with low vaccination rates, the precarious health conditions in host communities, and the congestion of health services in host communities. I use individual-level data from the Demographic and Health Survey (D.H.S.) from 1990 to 2017 to generate supporting evidence. First, I show no statistical differences in vaccination rates between host and non-host districts before and after 2007, with no differences between IDPs and native children. But, I observe a national immunisation rate decrease around 2007 which could increase the susceptibility to transmitting polio in overpopulated communities. Second, IDPs settle in communities with poorer conditions than non-host districts even before the large IDP inflow (i.e. higher number of members and children under five in the household, more likely to live in urban settings, and less likely to have a head of household working). Finally, I observe increased individual demand for health services in host districts after 2007. I measure the individual demand with the share of children with prenatal assistance.

To support the validity of my results, I estimate a dynamic difference-in-differences specification in which I calculate the difference in the polio cases between host and non-host districts on a yearly basis. The exercise supports the validity of the parallel trend assumption and shows similar impacts of IDP inflows to the ones identified in the aggregate regression. I also rule out the existence of confounders' effects from the conflict. For this purpose, I show that the conflict did not indirectly impact host communities' polio cases before the onset of the IDP crises in 2008, which validates my main findings. Additionally, I validate my treatment and counterfactual with alternative definitions. Finally, I test whether reverse causality between IDPs and polio exists.

The situation in Pakistan is not unique. In 2022, polio cases emerged in Malawi and Mozambique– two countries free of the virus for decades. Both countries are the scenario of a forced population movement from the conflicts ongoing in Northern Mozambique and the Eastern Democratic Republic of Congo, respectively. Protected children from diseases are far more likely to have the opportunity to thrive, the chance to learn and the ability to live healthy lives (UNICEF 2023). Therefore, three critical policy implications emerge from this paper. First, millions of forcibly displaced children migrate to camps or host communities. Since families in these settings are often transient, monitoring vaccination rates among these communities make it much harder to reach children with the necessary vaccines. Reaching the hard-to-reach- such as children from mobile and forced migrant populations or in conflict zones- should be a public priority (CDC 2021; UNICEF 2023). Second, poor communities are the host communities of most of the IDPs. An effort to better integrate the IDP population into the health services and labour market should be made to improve the conditions in which they live. Finally, the inflow of new population comes with increased demand for health services. Even if the increase in the demand is modest, in locations where the health delivery or capacity is weak, it can congest the local health services. It is essential to reinforce host communities' health workforce and infrastructure, so locals and newcomers can access health services equally.

This paper adds to the literature that studies the consequences of forcibly displaced populations in host communities (Becker and Ferrara 2019) by evaluating the impact on polio incidence, a disease close to being eradicated. As far as I know, the literature on the impacts of hosting displaced people on the spread of diseases is limited (Montalvo and Reynal-Querol 2007; Baez 2011; Ibáñez, Rozo, and Urbina 2021; Ibáñez, Moya, et al. 2023). None of the existing literature examines the effect of hosting internally displaced populations. This population does not cross an international border, making monitoring their health and vaccination status much harder. So, the impact of these inflows may differ from refugee inflows. Understanding the role of IDPs in polio-endemic countries is vital for formulating quick interventions in transit zones, camps and host communities. Additionally, I add to this literature by looking at an under-explored disease, polio, in an endemic country, Pakistan.

Second, my findings contribute to the research agenda on the determinants of infectious disease incidence. Most of the existing literature has focused on studying the mistrust of vaccines

(Martinez-Bravo and Stegmann 2022), the role of trade (Oster 2012), and public transportation closure (Adda 2016). In this paper, I study the role of internal displacement in spreading infectious diseases and explore some of the main transmitting mechanisms.

Finally, this paper belongs to the research agenda on the impacts of conflicts on health outcomes (Blattman and Miguel 2010; Devkota and Teijlingen 2010; Phadera 2021). I contribute to this literature by evaluating the impact of a direct consequence of conflict, forced migration, on the prevalence and transmission of polio.

2 Background

2.1 Conflict in F.A.T.A. Region

Pakistan witnessed a vast surge in violence after the terrorist attack in September 2001 in the United States (U.S.). The increase in violence manifested in waves of violent attacks against state institutions and civilians across Pakistan. Terrorists carried out around 1,600 attacks in the pre-9/11 era. However, a significant surge in the number of attacks was observed (around 12,000) in the aftermath of the 9/11 period (GTD 2021). The intensity of such violence was considerably higher in the Federally Administered Tribal Areas (F.A.T.A.) when the Tehrik-e-Taliban militants began entering into the region (Malik, Mirza, and Rehman 2023).[1]. Figure A.1 plots the number of attacks in Pakistan and the F.A.T.A. region.

F.A.T.A. was an autonomous tribal region in north-western Pakistan that existed from 1947 until being merged with the neighbouring province of Khyber Pakhtunkhwa in 2018.[2] F.A.T.A. were bordered by: Afghanistan to the north and west, Khyber Pakhtunkhwa to the east, and Balochistan to the south. Figure 1 illustrates the three administrative levels of Pakistan.[3] Its total population was estimated in 2000 to be about 3,341,080 people or roughly 2% of Pakistan's population, being Pakistan's most rural administrative unit. F.A.T.A. was located in *Pashtunistan* (*land of the Pashtuns* in Pashto), a historical pre-colonial region wherein Pashtun culture, the

[1]The Pakistani Taliban, formally called the Tehreek-e-Taliban-e-Pakistan, is an umbrella organisation of various Islamist armed militant groups operating along the Afghan–Pakistani border (Abbas 2008)

[2]The administrative units of Pakistan comprise four provinces, one federal territory, and two disputed territories: Punjab, Sindh, Khyber Pakhtunkhwa, and Balochistan; the Islamabad Capital Territory; and the administrative territories of Azad Jammu and Kashmir and Gilgit–Baltistan.

[3]A province (administrative level 1) has different divisions (administrative level 2), and a division is divided into other districts (administrative level 3).

Pashto language, and Pashtun identity have been based.[4]

The acceleration of violence in the F.A.T.A. led to a domestic and global policy response. After 9/11, Pakistani and U.S. forces exposed the F.A.T.A. to military offensives against alleged sanctuaries of terrorist outfits. On June 19, 2004, the U.S. undertook its first drone strike in Pakistan. Since then, the U.S. has carried out more than 406 drone attacks against alleged Al-Qaeda-linked affiliates in Pakistan's North-West. These attacks increased from 2007 and peaked around 2010. 98% of the drone attacks were in the F.A.T.A. Figure 2 shows the total number of drones from 2001 to 2022 (New-America 2021).

2.2 Forced Displacement within Pakistan

Since 2004, 13,289,880 million people have been displaced due to different operations against insurgents in F.A.T.A. Most of the affected population has been displaced multiple times after returning to their places of origin. 98% of the forcibly displaced population migrated within Pakistan. See Figure A.2 for a visual representation. The onset of the IDPs crises was in 2008, corresponding with a big jump in drone strikes in F.A.T.A. In 2009, the stock of internally displaced people (IDPs) reached more than 1.9 million individuals, corresponding to 57% of the F.A.T.A. population (UNHCR 2022). Figure 2 visually represents the total internally displaced population and the number of drones from 2001 to 2022.

The IDPs came from different F.A.T.A. districts, but especially from the most affected by the conflict (North Waziristan and South Waziristan) (UNHCR 2022). Figure 3 shows the positive correlation between the number of drones and IDPs from a given district. I present the total number of IDPs by origin in Table A.1.

Due to cultural and linguistic similarities, most IDPs migrated to relatively safe districts within the historical region of *Pashtunistan*. Many arrived in the place of displacement as cohesive groups, which helped them maintain a sense of community. They have also utilised social networks from their home areas (IDMC 2015). IDP's integration in host communities was not always easy. In specific locations, IDPs were even discriminated against by the native population and political leaders whipping up xenophobia against the displaced. Such attitude forced the displaced to find shelter only among their ethnic groups, adversely affecting their ability to relocate or freely seek

[4] *Pashtunistan* is a historical region on the Iranian Plateau, inhabited by the indigenous Pashtun people of southern Afghanistan and north-western Pakistan. During British rule in India in 1893, Mortimer Durand drew the Durand Line, fixing the limits of the spheres of influence between the Emirate of Afghanistan and British India and dividing the historical *Pashtunistan* as a share of two different countries (Bezhan 2014).

employment (Din 2010; IDMC 2015).

Additionally, IDPs usually reside in informal settlements in host communities and avoid living in camps for multiple reasons, including the fear of attack by non-state armed groups, poor conditions and lack of private space. Usually, the informal settlements lack safe drinking water, sanitation, and health care (IDMC 2015). I list the total IDPs in host districts by year in Table A.2.

2.3 Polio in Pakistan

Polio or Poliomyelitis is a highly infectious viral disease. The virus is transmitted by person-to-person contact. It lives in an infected person's throat and intestines. It spreads through contact with an infected person's stool (poop) or, less frequently, droplets from a sneeze. An infected person may spread the virus to others immediately before and up to two weeks after developing symptoms. The virus may live in an infected person's intestines for many weeks. They can contaminate food and water when they touch them with unwashed hands (CDC 2021).

There is no cure for polio, and it can only be prevented. Two polio vaccines are available: oral polio (OPV) and inactivated polio (IPV).[5] Children should usually get four doses of the polio vaccine at ages two months, four months, 6–18 months, and 4–6 years.

Although anyone not fully vaccinated against polio is at risk for polio, polio predominantly affects children under five. Most people who get infected with poliovirus will not have any visible symptoms. About 1 out of 4 people with poliovirus infection will have flu-like symptoms (fever, fatigue, headache, vomiting, stiffness of the neck and pain in the limbs). These symptoms usually last 2 to 5 days. One in 200 infections leads to irreversible paralysis (usually in the legs). Among those paralysed, 5–10% die when their breathing muscles become immobilised (WHO 2022).

Cases of polio have fallen dramatically over time. In 1988, more than 350,000 polio cases were reported annually across 125 countries. In 2021, the number of cases was down to 649. The main reason was the increased number of children vaccinated. Globally in 1980, only 22% of one-year-olds were vaccinated against polio, which increased to coverage of 86% of the world's one-year-olds in 2015.[6] In 2001, 14 countries reported cases of wild polioviruses. By 2021 there were only two countries where wild poliovirus cases were recorded: Afghanistan and Pakistan (WHO 2022).

[5]OPV is administered orally and can be given by volunteers. OPV protects both the individual and the community because it induces gut immunity, which is essential to stopping poliovirus transmission. IPV is given by injection and needs to be administered by a trained health worker. IPV is highly effective in protecting individuals from severe diseases caused by poliovirus. However, it cannot stop the virus's spread in a community.

[6]In 1988, the World Health Assembly created the Global Polio Eradication Initiative to eradicate polio by 2000.

Since 1994, the Pakistan Polio Eradication Programme (P.P.E.P.) has been fighting to end the crippling poliovirus in the country. In 1997, Pakistan reported 1,147 cases, constituting 22% of the cases reported worldwide. With its initial extraordinary efforts to control polio among children, Pakistan reduced cases from 20,000 in 1990 to 28 in 2005. However, about 100 cases have been reported annually after 2007. Cases steadily rose from 32 in 2007 to 118 in 2008 to 198 in 2011.

2.4 Inmunization in Pakistan

Children in Pakistan typically receive three primary vaccines through routine immunisation activities: the vaccine against poliomyelitis, the D.P.T. (vaccine against diphtheria, pertussis, and tetanus) vaccine, and the measles vaccine. Pakistan follows the recommended vaccination calendar of the World Health Organization, and the first dose of most vaccines is supposed to be administered shortly after birth.

As part of the P.P.E.P., Lady Health Workers are the health workers responsible for child immunisation. These workers are assigned to a local health facility, each responsible for approximately 1,000 people or 150 homes. They regularly visit households to provide family planning information and immunise children according to the vaccination schedule. Since 2010, the provision of public health goods is a provincial responsibility. In 2014, there were approximately 110,000 Lady Health Workers in Pakistan. However, the main way Pakistani children are immunised is through vaccination drives. There are national and subnational immunisation days during which vaccinators (typically lady health workers joined by other volunteers) provide vaccines at households' doorstep. They typically last for three days and target all children up to age 5 in the respective district. All the vaccines provided during immunisation drives or at public health facilities are free of charge (Martinez-Bravo and Stegmann 2022).

The surge in violence in F.A.T.A. could be one of the leading reasons behind the increase in polio cases in Pakistan. Almost 70% of Pakistan's polio cases from 2004 to 2018 were reported from this area. Unhygienic and poor sanitary conditions with large families living in packed houses resulted in widespread polio transmission. Moreover, the militants carried out continuous propaganda against polio vaccination, translating into increased vaccine refusal. As the extremists banned polio vaccination, almost 400,000 children could not be vaccinated in the tribal north during 2010–2011. Even vaccination workers began to be attacked and killed (Mushtaq et al. 2015; Rahim, Ahmad, and Abdul-Ghafar 2022).

The movement of the population during the conflict has led the P.P.E.P. to implement a particular program targeting the High-Risk Mobile Populations, or H.R.M.P.s (nomads, Internally Displaced Persons, Afghans, brick kiln workers and visiting "guest children"). The H.R.M.P. strategy requires vaccinating all eligible children at all possible opportunities, including in departing communities, transit, and communities where they settle. The P.P.E.P. vaccinates children travelling or on the move through 500 permanent transit points (P.T.P.s) across all major transit points nationwide. These P.T.P.s are set up along country and district borders and other essential transit points such as railway stations, bus stops, and highways. In 2018, P.T.P.s had vaccinated a total of 1.7 million children (UNICEF 2019).

3 Data

I construct a panel dataset at the district and monthly level that combines data on conflict, total forcibly displaced population, polio cases and supply-demand of vaccines.

3.1 Conflict data

I use two georeferenced variables to measure conflict intensity in Pakistan—the number of drone strikes and terrorist attacks at district and monthly levels.

The conflict data on drone strikes comes from the World of Drones Database developed by New America (New-America 2021). New America gathers information on each drone strike's timing (day, month, year), location (latitude and longitude) and total deaths. The World of Drones database draws upon media reports and other open-source information to track which countries and non-state actors have armed drones or are developing them; and which actors have used them in combat and where.

The New America Database has reported 406 drone strikes in Pakistan from January 1, 2001, to December 31, 2022. The first drone was recorded on June 19, 2004, and the last on July 4, 2018. Only 10 of the 406 drones were located outside F.A.T.A. Figure 3 presents the spatial distribution of drone strikes in Pakistan. I construct my primary measure of conflict by aggregating the drones that fall in a district monthly. To show the robustness of my results to alternative ways of measuring drone intensity, I construct a supplemental measure: the number of people killed. Figure A.3 presents the spatial distribution of total deaths by drone strikes in Pakistan.

The data on attacks against the state and civilians are extracted from the Global Terrorism Database - G.T.D. (GTD 2021). The G.T.D. provides details on more than 200,000 terrorist incidents worldwide since 1970. For each incident, information is provided on the timing (day, month, and year), location (latitude and longitude), fatalities (wounded and killed), type (assassination, explosion, suicide, hijacking, etc.), target (civilians, businesses, government officials, religious institutions, N.G.O.s, etc.), the terrorist group which carried out the attack, and the motivation of the episode (political or religious).

The G.T.D. reported 13,638 terrorist attacks from January 1, 2001, to December 31, 2020. I construct a measure of terrorist attacks at the district level by aggregating the number of incidents that fall in a district. I complement this measure by repeating this exercise with the number of people killed in the attacks.

3.2 Forced displacement data

United Nations High Commissioner for Refugees - U.N.H.C.R. provides the data on forcibly displaced populations (UNHCR 2022). UNHCR 2022 contains information about the countries of destination and origin, province and district within a country, total population, year of arrival, and demographic characteristics (age and gender). Therefore, this data allows me to identify the total internally displaced population (IDPs), total Pakistani refugees outside Pakistan, and total Afghan refugees in Pakistan. Figure A.2 shows how a large share of the forcibly displaced population remained within Pakistan.

Among the IDPs who fled from F.A.T.A., 54% of them are below 18 and 18% below 5. Figure A.4 plots the total IDP distribution by age. Moreover, 47% of the IDPs were women or girls. The destination districts were concentrated in Khyber Pakhtunkhwa, as shown in Figure A.5. Figure 1 shows the spatial distribution of the host districts reported by U.N.H.C.R., which all fall within the historical *Pashtunistan*. Figure A.6 plots the total IDPs in each district over time.

3.3 Polio data

I collect data on polio incidence from January 1, 2001, to December 31, 2022, from the Polio Eradication Program established by the World Health Organization (WHO). The Polio Eradication Program gathers information for each reported polio case on the timing (year, month and year), location (district), and the type of virus. I build my outcome measure on polio incidence by aggregating the number of new polio cases in a given district and month.

The Polio Eradication Program reported 2,080 new polio cases from 2001 to 2022. The cases in the entire country and F.AT.A. have followed a similar pattern. Figure A.7 shows the evolution of cases. There are three critical years where the trend switched to positive: 2008, 2012 and 2018.

3.4 Vaccination supply and demand

For this project, I also collect data on Pakistan's polio vaccination campaigns between 2001 and 2022. I obtain this data from the Polio Eradication Program. These data contain district-month measures of whether a polio vaccination campaign was conducted, the type of campaign—case response, mop-ups, child health days, subnational or national immunization days—, the age group targeted, and vaccine type.[7]

I rely on data from two waves of the Demographic Health Surveys (D.H.S.) in Pakistan to obtain measures of polio immunization at the individual level from 2008. In particular, I used information on the demand for the polio vaccine from the 2012/13 and 2017/18 D.H.S. surveys. Moreover, I profit from the 2006/07 and 1990/91 D.H.S. surveys to obtain measures before 2008. The D.H.S. has data on the year and month of birth, allowing me to define the exposure to the inflow of the IDP population.

The D.H.S. asks each household member whether the individual was born in the current district of residence and the reason for the migration. I exploit this migration data to build a variable on whether an individual is displaced or native in a given district. The D.H.S. also contains georeferenced household location information (only in 2006/07 and 2017/18 D.H.S. surveys).

Finally, the D.H.S. characterize the host communities at the household (i.e. sanitation, overcrowding, house conditions, and health provision) and individual level (i.e. health-seeking behaviour, labour, and education).

3.5 Other data

For my empirical strategy, I identify the historical pre-colonial region of *Pashtunistan* in Pakistan from the Georeferencing Ethnic Power Relations - GeoEPR 2021 dataset (Vogt et al. 2015). GeoEPR geo-codes all politically relevant ethnic groups from the Ethnic Power Relations-Core 2021 dataset and provides polygons describing their location on a digital map.

[7] Mop-ups are very targeted, geographically limited polio campaigns, held between larger-scale national Immunization Days or subnational Immunization Days in areas where we know many children were missing, for example, or where a large immunity gap persists. Child Health Days are not specifically polio campaigns, but the polio vaccine is added to Child Health Days alongside other vaccines and health interventions. National Immunization Days are nationwide campaigns targeting all children aged 0-5. Subnational Immunization Days are vaccination campaigns in key high-risk provinces.

To test for the validity of my identification strategy, I use additional controls, including constructed district-year level data on satellite night light density as a proxy of economic development. The National Oceanic and Atmospheric Administration (NOAA) processes night light density data. NOAA uses satellite images collected by the U.S. Air Force Defense Meteorological Satellite Program. Two satellites that circle the Earth 14 times daily collect the images, recording the intensity of Earth-based lights with their Operational Linescan System. I also use sociodemographic data from the 1998, 1981, and 1973 Population Census. [8]

4 Identification Strategy

This paper aims to study the impacts of inflows of internally displaced populations on the hosting communities' polio incidence. To tackle this question, my identification strategy relies on comparing district new polio cases in locations exposed to large IDPs inflows with new polio cases in those minor or non-affected before and after the onset of the IDP crises of 2008.

IDPs settlement is a potential endogenous decision, and time-varying characteristics in host communities could affect the resettlement pattern and polio incidence. For example, IDPs might choose to migrate to poor areas closer to their original communities, leading us to overestimate the harmful effects of IDP population on the number of polio cases. Additionally, most IDPs settle down in host communities without registration systems (UNHCR 2017). Herefore, the officially identified host districts and the total IDPs in each community may be underestimated. To overcome these challenges, I exploit the proximity to the *Pashtunistan* historical border to define the treatment and counterfactual.

4.1 Pashtunistan historical border

F.A.T.A. is part of the pre-colonial region of *Pashtunistan*. And, due to cultural and linguistic similarities, many of the IDP families from F.A.T.A. migrated to other districts within *Pashtunistan*, as shown in Figure 1. The border between Afghanistan and Pakistan (known as *the Durand Line*) results from an agreement in 1893 between the British Indian government and the emir of Afghanistan. Eighty-five per cent of the Durand Line follows rivers and other physical

[8] The 1998 Census sample covers a share of Khyber Pakhtunkhwa (23 out of the 31 districts) and Punjab (24 out of the 38 districts). The 1973 Census sample covers all Balochistan (28 districts) and Punjab (38 districts), a shared of Khyber Pakhtunkhwa (18 out of the 31 districts) and Sind (26 out of the 27 districts). The 1981 Census has information for 76 out of the 141 districts in Pakistan.

features, not ethnic boundaries, splitting the historical Pashtu region into two separate countries.

The native people of *Pashtunistan* are the Pashtuns. They are the largest ethnic group in Afghanistan and the second largest in Pakistan. The main language spoken in the delineated *Pashtunistan* region is Pashto. The Pashtuns practice Pashtunwali, the indigenous culture of the Pashtuns. This pre-Islamic identity remains significant for many Pashtuns and is one factor that has kept the *Pashtunistan* culture alive. Although the Pashtuns are politically separated by the Durand Line and other administrative borders within Pakistan, Pashtun tribes tend to ignore the borders. For instance, many Pashtun tribes from the F.A.T.A area and the adjacent regions of Afghanistan used to cross back and forth with relative ease to attend weddings and other events. After 2004, this cross-border movement is checked via the military and has become much less common compared to the past. However, the transit across *Pashtunistan* districts in Pakistan has never stopped, allowing IDPs to move within their historical region.

To identify the effects, I compare the polio cases in districts within *Pashtunistan* (treatment) with those in districts immediately outside *Pashtunistan* (counterfactual) before and after 2007. Figure 4 shows the Pakistani districts within and outside Pashtunistan. The central identifying assumption is that the IDP population mostly moved to Pashtu districts, and non-Pashtu districts had no or negligible presence of IDPs. Nevertheless, there are two features worthy of highlighting. First, there are some districts in which only a share of their territory is within *Pashtunistan*. I start by removing them from the sample since I can not define them as being within or outside *Pashtunistan*. Second, *Pashtunistan* covers F.A.T.A., but I drop it from the sample to avoid potential confounding conflict effects. Hence, the treated districts are those on the left side of the *Pashtunistan* border and the control districts on the right. The design of my counterfactual allows me to compare host districts (Pashtu districts) to the most similar administrative units possible (the closest non-Pashtu districts). I use alternative definitions of treated and control districts in the robustness section. See Figure 4 for a spatial distribution of the treatment and counterfactual.

I estimate the following specification:

$$Y_{d,p,t,m} = \beta_1 Pashtunistan_d \;\; X \;\; IDP \;\; Crises_t + \beta_2 X_d + \alpha_p + \gamma_{t,m} + \epsilon_p \tag{1}$$

where d stands for district, p stands for province, t for year and m for month. $Y_{d,p,t,m}$ represents the district outcome: one if at least a polio case in the year and month t-m, zero otherwise. $Pashtunistan_d$ stands for being or not within *Pashtunistan*: one if district d falls in the historical region, zero otherwise. $IDP \;\; Crises_t$ is a dummy variable that takes the value of one after 2007. X_d is a matrix of district-year controls. Namely, I control for district-level nightlight intensity as a proxy of economic development (Pérez-Sindín, Chen, and Prishchepov 2021) and the total number of polio vaccination campaigns in a given district month which account for the vaccination supply. α_p and $\gamma_{t,m}$ account for province and year-month fixed effects. Standard errors are clustered at the province level to account for time serial correlation in the outcome across geographic areas. A battery of robustness tests that support the validity of my identification strategy is presented in sections 5.2 and 7.

5 Results

I first examine whether there are different likelihoods of having at least one polio case in IDP population host districts to districts with no or negligible presence of IDP. The specification estimates in equation (1) are presented in Panel A of Table 1.

The results suggest that districts that received the IDP population have experienced an increase in the probability of polio incidence compared to non-host districts. Column 1 shows the estimates without fixed effects and controls. An IDP inflow increased the probability of at least one polio case in the host districts by 3.8 percentage points (pp). The estimates are statistically significant at the one per cent level. The magnitude of the effects does not change when adding province-fixed effects and year-month fixed effects (column (2)), but the estimates are significant at the five per cent level. Column (3) shows that the point estimates increase to 5.3 pp when controlling for nightlight intensity and total vaccination campaigns (significant at the five per cent level). We could be concerned that the different characteristics in sanitation and overcrowding between treated and control districts could drive my results.

In column (4), I control for the average number of children under five, the average number of members in a household, and the total share of the literate population in a district from the 1973, 1981 and 1998 Population Census. The results hold and are significant at the 10 per cent level. The issue with the population census is that none of the censuses covers the entire sample. So, to increase the coverage, I combine them. However, it implies that the covariates are measured at different points in time. In column (5), I control for contemporary characteristics instead (the average number of children under five, the average number of members in a household, shared households with piped water, and shared households with a finished floor). The magnitude of the effects slightly increases and is significant at the five per cent level. When I look at the absolute number of polio cases in Panel B of Table 1, the results hold.

5.1 Contextualizing the Magnitude of the Effects

What is the impact on polio cases relative to the population size? I obtain similar results when estimating equation (1) with a continuous outcome: the number of new polio cases per 100,000 inhabitants. I present the results in panel C of Table 1. I use the population in 2017 to calculate the number of cases per 100,000 inhabitants.[9] An inflow of the IDP population results in 0.007 additional cases per 100,000 inhabitants (column (3)). Although negligible, this represents an increase of 40% over the mean incidence.[10] The intuition behind this finding is that host districts are more likely to encounter a polio case and to experience a more rapid spread of the virus.

Additionally, I use the population in 1998 to estimate the polio cases per 100,000 native inhabitants. Table A.3 shows that an inflow of the IDP population corresponds with 0.012 additional cases per 100,000 native inhabitants (see column (3)). The points estimates are statistically significant at the five per cent level.

Does the intensity of the IDP inflow affect the results? Precise data on the IDP inflow at the district-year level does not exist for my entire timeframe. Therefore, I approximate district-year inflows of the IDP population using the following measure:

$$PredictedInflow_{dpt} = IDPInflow_t \text{ x } \frac{1}{distance_{dp}}$$

where $IDPInflow_t$ represents the total inflows of the internally displaced population registered in Pakistan in each year t, and $distance_{dp}$ is the Euclidean distance from the centroid of each district d from the province p to F.A.T.A. border. I construct my predicted inflows measure as the interaction of the inverse distance of each district to the closest F.A.T.A. border (district variation) and the total yearly number of IDP population (annual variation).

The distribution of $PredictedInflow_{dpt}$ across the inverse distance of each district to the F.A.T.A. border is displayed in Figure A.8. To test whether this cross-section variation is associated with IDP migration patterns, I use the available data on the IDP population at the district level from the U.N.H.C.R. Figure A.9 shows that my inverse distance measure correlates highly with the reported number of IDP populations by U.N.H.C.R.

Panel A in Table 2 shows the equation (1) results with continuous treatment, $PredictedInflow_{dpt}$. The intensity of the IDP inflow has a significant effect on polio incidence. An increase of one standard deviation in predicted inflows results in 0.001 additional cases per 100,000 inhabitants (see

[9] Azad Jammu and Kashmir and Gilgit-Baltistan provinces are not in the 2017 Population census. So, Kargil, Kupwara, Muzaffarabad and Neelum districts drop from the sample.

[10] The average number of new polio cases in my sample from 2001 to 2022 equals 0.005.

Panel C in Table 2), corresponding to a 20% of the mean incidence.

Why are the main effects meaningful? Eradicating polio has been a worldwide effort over the last decades. But, until polio is completely eradicated, all countries remain at risk of imported wild poliovirus. Identifying the determinants of new cases is critical to prevent additional ones. Polio, which only infects humans, has been eliminated in 193 countries. With the transmission of wild-type polio limited to Afghanistan and Pakistan, an official eradication declaration is in sight. In 2005, 28 cases were reported in Pakistan, compared to the 1,147 cases in 1997. Moreover, most host districts had zero or close to zero polio cases before 2007. Ultimately, the estimates I present in this paper capture the impacts of IDP inflows on keeping host communities away from eradicating polio rather than the effects of an extensive increase in polio cases. Even if these results are substantial in magnitude. A high risk of underreporting is present, as 75% of people infected with poliovirus are asymptomatic (WHO 2022). Ideally, I could look at the incidence of other diseases, such as measles, chickenpox, or malaria, to validate my findings. Unfortunately, I could not find comparable data for other diseases.[11]

5.2 Threats

Unbalanced pre-treatment characteristics. Table A.4 presents summary statistics of baseline key demographics and socioeconomic variables that compare districts with high IDP intensity ($Pashtunistan_d$ = 1) and districts without or with negligible IDP population ($Pashtunistan_d$ = 0). The balancing tests do not reveal significant pre-shock differences except that treated districts seem to have a higher share of households with piped water, toilets, television, and in urban locations, and an additional number of children under five and four more members in the household on average terms. However, by definition, the regions may have differences in social beliefs and norms. Since in host districts, Pashtu traditions and norms are predominant in host districts. Table A.16, in the robustness section, shows similar results with a counterfactual within the pre-colonial region of *Pashtunistan*, ruling out the possibility that social norms differences could bias the main findings.

Pre-trends outcomes. The findings indicate that there are no pre-trends in the outcomes. The key identifying assumption of the main results is that host and non-host districts should evolve similarly without treatment. In other words, there may not exist differences in the pre-treatment

[11] The information on new polio cases from 2001 to 2022 is not publicly available. I obtained this information from the Expanded Program of Immunization. The data available for the other diseases were insufficient to grant a meaningful regression analysis or were not available before 2008.

outcome between treated and control districts. Figure 5 plots the new polio cases from 2000 to 2022 in treated and host districts, suggesting that non-consistent differences exist. I generate additional evidence in Figure A.10, estimating an event study with the new cases per 100,000 inhabitants.

Conflict effect. Conflict can affect the health outcomes of children at early age (Bundervoet, Verwimp, and Akresh 2009). The conflict is primarily concentrated in F.A.T.A., which is not in my baseline sample. Still, we could be concerned about potential spillovers in the neighbouring district to F.A.T.A., which could bias my findings. The 86% of the drone strikes were located on the southern divisions of F.A.T.A. (in North Waziristan, South Waziristan and Bhittani districts). We could assume that adjacent districts to F.A.T.A.'s northern districts were less indirectly exposed to the conflict. Table A.5 shows in panels A and B that the results of Table 1 hold when I restrict my treatment and counterfactual to the neighbouring northern or southern districts.[12]

Terrorist attacks took place across Pakistani districts. In particular, 43% of the districts of my sample experienced at least a terrorist attack during my period of analysis (2001 to 2022). The intensity of the attacks could contaminate my estimates. I control the number of attacks in a given district to rule out this hypothesis. The points estimates remain statistically significant, but the point estimates slightly decrease to 0.044 (see panel C of Table A.5).

Afghan refugees. Since the late 1970s, Pakistan has been a host country for millions of refugees and some 1.35 million still reside in the country. (UNHCR 2022). Figure A.11 shows the evolution of total Afghan refugees in Pakistan from 2001 to 2022. Most refugees are in the Pashtun-dominated areas of Pakistan. This fact is a major problem for my identification. To upfront this empirical limitation, I conduct three different exercises. First, I show that the results of Table 1 hold when I control for the total district-year refugees (see panel A of Table A.6). Only 3 out of the 296 camps had IDPs as the targeted population. It made that most of the refugees live in refugee camps. Thus, as a second exercise, I also show evidence that the estimates do not change when I control for the number of camps in a district. Finally, the results do not substantially change when I add an interaction to the number of camps in a district. See the results in panels B and C of Table A.6. The 67% of the treated districts have a refugee camp compared to the 26% of the non-host districts. So, even if all these checks hold, I can not rule out

[12] The districts in the North are: Abbottabad, Attok, Chakwal, Charsadda, Hangu, Islamabad, Kargil, Kupwara (Gilgit Wazarat), Malakand P.A., Muzaffarabad, Neelum, Peshawar, and Rawalpindi. Panel B shows the estimates for a sample of Southern districts (Bannu, Bhakkar, Bhittani, Bolan, Chagai, Dera Bugti, Kalat, Kashmore, Khushab, Layyah, Musakhel, Muzaffargarh, Pishin, Qilla Saifullah, Rajan Pur, Tank, Zhob, and Ziarat).

completely that my estimates do not capture the effects of refugees itself.

Migration outflows. Although very few Pakistanis migrated internationally, a big jump in the number of Pakistani refugees before and after 2007 could affect my results. Figure A.2 helps to remove this concern. The number of Pakistani refugees has been relatively constant from 2000 to 2011, with an increase from 2012. However, the results remain unchanged when I restrict the time horizon of my analysis until 2011 (see Table A.7).

6 Mechanisms

There are three channels through which IDP inflows could increase polio incidence in host communities. First, low vaccination rates at the national level increase the susceptibility to transmitting polio in overpopulated communities. Second, the precarious health conditions in host communities may facilitate the spread of the virus. Third, a sudden increase in the population could congest the health services in host communities. Yet, one caveat of the identification strategy implemented in this paper is that it cannot disentangle the precise mediating mechanisms underlying the observed increase in the number of new polio cases reported in treated communities with respect to control communities. Notwithstanding this shortcoming, in this section, I provide empirical evidence below that, although they are not conclusive, the evidence suggests that the three proposed mechanisms could have a role in the increase in polio incidence. I use individual-level data from the Demographic and Health Survey from 1990 to 2017 to analyze the mechanisms.

6.1 Overpopulated communities with low inmunization rates

Pakistan is one of the most populous countries in the world and one of the least developed, with a large population of approximately 188.9 million people, including 24.7 million children under five years old (UNDESA 2015). Pakistan's sociodemographic characteristics make polio eradication critical. Conflict and insecurity affecting routine immunisation teams' access generated a decline in immunisation coverage to less than 45% in F.A.T.A. (Hussain et al. 2016). As a result, we could expect that the massive arrival of the population from the F.A.T.A. decreased the immunisation rates in the host communities, increasing the number of new polio cases. To test this hypothesis, I use individual-level information on vaccination against polio and the date of birth from the Demographic and Health Survey (D.H.S.). First, Figure A.12 shows no differences in the share of children vaccinated against polio with at least one dose between treated and control districts after

2007. Second, I estimate the following specification to exploit within-district cohort variation:

$$Y_{i,d,k} = \beta_1 Pashtunistan_d \ \ X \ \ IDP \ \ Crises_k + \beta_2 X_i + \alpha_d + \gamma_k + \epsilon_d \quad (2)$$

where $Y_{i,d,k}$ is equal to one if child i from the cohort k living in district d received at least one dose of the polio vaccine, zero otherwise. The timing of the treatment is given by the year and month of birth: $Crises_k$. $Crises_k$ is one if child i was born after December 2007. I control for timing covariates at the district and individual levels. District-level covariates include nightlight intensity and the number of polio activities in the year of the interview. And I include the work status of the head of household, urban location, and gender of the child as individual-level covariates. Additionally, I include district α_d and cohort γ_k fixed effect, which accounts for generational changes in immunisation supply or social patterns.

I find that children in host communities born after the arrival of the IDP population are less likely to be vaccinated than those born before. Panel A of Table 3 shows the results. It seems that the IDP inflow decreases the probability of immunisation in the host districts by 6.4 percentage points (pp) in Column (2). The estimates are statistically significant at the ten per cent level. The magnitude of the effects does not change with covariates. Column (3) shows that the point estimates increase to 8 pp when controlling for nightlight intensity and the total vaccination campaigns in a year (significant at the ten per cent level). We could be concerned that the employment or sociodemographic characteristics could drive my results. In column (4), I control for the head of household's work status, urban location, and gender of the child. The results hold and are significant at the five per cent level. The results also hold when I simultaneously control for the covariates of columns (3) and (4) (see column (5)). These findings suggest that decreasing immunisation rates over time could have affected the number of polio cases.

One caveat of the above-results is that I can not disentangle if the inflow of the IDP population drives the decrease in polio immunisation in host districts. Similar changes may happen in the non-host districts. Therefore, to account for the spatial variation between host and non-host communities, I now change the district-fixed effects of equation (2) for province-fixed effects. For this new specification, there are no differences between treated and host districts in the share of children vaccinated after 2007 (see Table A.8). Figure A.12 supports this finding. We can observe how the share of children vaccinated is pretty similar between host and non-host districts over time, with a decrease around 2007.

Then, how could immunization rates affect the results? In some instances, the population of villages and towns doubled within a brief timeframe with the influx of the IDP population. The vaccination rates after 2007 barely reached 40% in both treated and control districts. Hence, such low immunisation rates in overpopulated host communities facilitate polio transmission. This hypothesis is very in line with the increase in polio incidance or Table 1.

Are the IDP children less likely to be vaccinated? I generate three pieces of evidence suggesting no statistical differences between IDP and native children. First, I add an interaction dummy in equation (2), being one if child i is internally displaced, 0 otherwise. Panel B of Table 3 shows that the interaction is not statistically significant. In the 2012 and 2017 waves of the D.H.S., there is a self-reported question on the reasons to migrate with an answer on "escape war or violence". I use this information to build my dummy variable. However, we need to assume a certain level of underreporting, which could affect the reliability of the results. Second, we would expect that if IDP children are less likely to be vaccinated, a higher intensity in the number of IDPs higher the aggregate number of non-vaccinated children. In Panel C of Table 3, I repeat the exercise of Panel A with the predicted inflow. The points estimates are statistically non-significant. Third, I conduct the same analysis of panel A of Table A.8, including province-fixed effects and the predicted inflow in panel B. The results are non-significant, either.

Although the vaccination of children decreased strongly in F.A.T.A. from the beginning of the conflict, a vaccination programme targetting children on the move could be the main reason behind the above results. For Pakistan's polio programme, the High-Risk Mobile Populations, or H.R.M.P.s– nomads, Internally Displaced People, Afghans, brick kiln workers and visiting "guest children"– are critical. The H.R.M.P. strategy requires vaccinating all eligible children at all possible opportunities, including in departing communities, transit, and communities where they settle. The Pakistan Polio Eradication Programme vaccinates children travelling or on the move through 500 permanent transit points (P.T.P.s) across all major transit points nationwide. These P.T.P.s are set up along country and district borders and other essential transit points such as railway stations, bus stops, and highways. The programme has developed impressive strength for vaccinating H.R.M.P. on the move. In 2018, P.T.P.s had vaccinated a total of 1.7 million children. For instance, the National Emergency Operations Centres of Pakistan have vaccinated children under ten at major transit points-border areas in southern Khyber Pakhtunkhwa, as this area has been home to some of the more high-risk mobile populations. Figure A.13 shows the location of the P.T.P.s surrounding the south of the F.A.T.A. region (UNICEF 2019).

In sum, one major channel of how an IDP inflow could affect the results may be the sudden increase in the population size in host communities. In a country with low immunisation rates, overpopulated communities could become suitable locations for new polio cases.

6.2 Poor Conditions in Host Communities

Many IDP families migrated to informal settlements, Pashtun slums or were squeezed into the houses of friends or relatives. Access to safe drinking water and hygiene is a significant problem for them. Appropriate facilities for bathing, doing laundry or keeping personal hygiene are not always available, facilitating the transmission of polio (IDMC 2015).

One crucial question is whether IDP settle in poorer locations or if the living conditions get worst with the sudden arrival of new population. The results of Table A.9 suggest that IDP population move to the poorest locations. Table A.9 shows how the number of household members and children under five was larger in host districts than in non-host districts before 2008. Moreover, Table A.9 presents evidence that households in host districts were also more likely to live in urban settings and less likely to have a head of household working before 2008. These pre-treatment characteristics may be a key channel behind the main results and, as well, a vital identification threat. Even if I control for local economic development, I can not ensure that my estimates capture the actual impact of IDP inflows rather than the pre-treatment differences in disadvantages characteristics. Nonetheless, what is certain is that poorer communities cannot respond efficiently to an IDP inflow, which implies that they are systematically more affected by the waves of displaced persons. Keeping this limitation in mind, how does the arrival of IDPs affect the local health conditions? To shed light on this question, I estimate equation (3) on six outcomes related to household conditions in host communities (i.e. access to drinkable water, access to a toilet, floor quality, number of children under five, households member, and head of the household's working status).

$$Y_{h,d,t} = \beta_1 Pashtunistan_d \ X \ IDP \ Crises_t + \beta_2 X_h + \alpha_d + \gamma_t + \epsilon_d \tag{3}$$

where $Y_{i,d,t}$ is equal to one if household h living in district d has a given household characteristic at the time of the survey t. The timing of the treatment is given by the year of the survey: $Crises_t$. $Crises_t$ is one if household h was interviewd after December 2007. I control for nightlight intensity and urban location. Additionally, I include district α_d and time of survey γ_t fixed effect, which accounts for time-invariant covariates across households in a district.

I find that households exposed to the IDP inflow decrease the probability of having a piped water system compared to non-exposed households (see Panel A of Table 4 for the results). However, I do not observe an increase in cases of diarrhoea or fever. However, that is not the case (see Table A.10). I obtain the same results with the predicted inflow measure (see panel B of Table 4).

In panel C, I look at the heterogeneity between IDP and native families. IDP households are less likely to have additional children and a head of household working but more likely to have an additional household member and piped water.

The above results suggest that IDPs settle in communities with poorer conditions where a new polio case can quickly multiply. Still, the quality of the communities deteriorates with the arrival of the IDP population, with the IDPs as the most affected.

6.3 Congested Health Services

An alternative hypothesis could be that the displaced families created logistical hurdles in delivering subsistence and healthcare assistance to the scattered communities. Furthermore, the increased demand for healthcare services could have caused additional strain on the local infrastructure, which was often hardly adequate even for the needs of the local population (Din 2010). As a result, the arrival of new populations (IDP families) may have restricted access to quality health services for native and IDP populations, affecting the incidence of polio.

Conceptually, by capturing the aggregate demand in health services, I would be able to shed some light on the potential changes created by IDP inflows. Unfortunately, this information is not available. I can only capture the individual demand by using individual-level data on prenatal and postnatal doctor assistance from the DHS. Figure A.17 and Figure A.18 illustrate an increase in the share of children with prenatal assistance and a slight decrease in postnatal assistance after 2007 in host districts. However, I observe a decrease in prenatal and a more substantial decrease in postnatal services in non-host communities. To complement the descriptive analysis, I repeat equation (2) with the prenatal and postnatal doctor assistance outcomes (columns (1) and (4)). I include province-fixed effects in columns (2), (3), (5), and (6). Results in Table 5 (panel A) do not support the idea of changes in the individual demand for health services in treated districts after the IDP inflow. This result is unsurprising since it is hard to imagine that the IDP inflow generates a behavioural change in the demand for health services.

Is the supply responsive to an increase in the demand for health services? Ideally, I would like to study this question using district-level data on health service delivery (health centres and workforce). Unfortunately, I could not get this data. To address this limitation, I proxy health services supply with district-level data on polio vaccination campaigns from the Polio Eradication Program. In Pakistan, health is primarily the responsibility of the provincial government. Therefore, the central assumption is that the supply of health services follows the same pattern as the polio vaccination campaigns. Figure A.14 shows how the total vaccination activities increased in 2008 in host districts with respect 2007. I observe a similar increase in the total vaccination activities per 100.000 inhabitants (see Figure A.15).[13] Results of Table 5 support the idea of a responsive supply—an increase of one standard deviation in predicted inflows results in 0.06 additional vaccination campaign. However, I can not disentangle if the increase in the supply is high enough to meet the demand for formal health services.

6.4 Other mechanisms

After the conflict in F.A.T.A., the resistance to foreign interventions has considerably increased across the country. Along these lines, the misinformation and suspicion regarding the polio vaccination have also been a barrier to stopping polio eradication in Pakistan. These concerns include misconceptions regarding the vaccine's efficacy in local communities and among vaccinators engaged in repeat polio campaigns. Among the rumours, there is a widespread belief in Pakistan and elsewhere that the vaccine causes girls' infertility, impacting the reduction of vaccinated girls (Rahim, Ahmad, and Abdul-Ghafar 2022). The disclosure of the C.I.A. vaccination ruse could have significantly impacted mistrust of vaccines and "Western" health services because it lent credibility to many of the Taliban's arguments against vaccines.[14] The mistrust in health services could be a potential underlying channel behind the increase in polio cases. Vaccine refusal has been mostly restricted to Khyber Pakhtunkhwa and F.A.T.A. Therefore, the IDP population from F.A.T.A. could transmit the misinformation to host communities. I repeat equation (1) to evaluate this hypothesis and restrict my time horizon until June 2011. Table A.11 present significant estimates when restricting my time horizon. Additionally, the results of Table 3 on immunization do not change when I drop the months after June 2011 (see panel B of Table A.11). Due to the infertility rumours, girls were the most affected by the vaccine refusal. Panel C of Table A.11 shows no statistical differences between host and no-host districts in the vaccination rates of girls. These findings help rule out the hypothesis that the IDPs could be active in misinformation transmission.

Finally, larger IDP inflows may increase prices through higher local demand or displace informal

[13] I use the aggregate population stock in host and non-host districts using the 2017 Population Census. When I use the aggregate population using the 1998 Population Census, I obtain a similar graph in Figure A.16.

[14] The C.I.A. wanted to obtain definite proof that Bin Laden was hiding in Abbottabad, Pakistan. To this end, the C.I.A. organised a vaccination ruse. The objective was to obtain D.N.A. samples of children living in the compound and compare them to the D.N.A. of Bin Laden's sister, who had died in Boston in 2010. On July 11th of 2011, the British newspaper *The Guardian* published an article describing the vaccine ruse (Martinez-Bravo and Stegmann 2022).

workers from the labour markets, thus deteriorating health outcomes through worse economic conditions. According to the Internal Monitoring Center, IDPs also face hostilities from the host communities. The perception exists that IDP workers substituted native workers because they accept lower pay than the local population and force wages down. Table 4 shows how the head of households in IDP families is less likely to have jobs than natives. But, this information does not allow me to examine the local labour market dynamics. So, given data constraints, I can not test this mechanism empirically.

7 Robustness Checks

I present evidence of the validity of my results in three ways. First, I conduct two falsification tests to rule out hidden effects. Second, I test the validity of my treatment and counterfactual using alternative definitions. Third, I approach my research question using alternative specifications.

7.1 Falsification Tests

In this project, I look at the impacts of hosting conflict-induced IDPs on polio incidence in host communities. A major concern is that the effects of Table 1 could be driven by the effect of conflict rather than by the IDP inflow. We should observe no effect on host districts before the treatment to reject this hypothesis. In this setting, the violence surged after the terrorist attack in September 2001. However, the mass movement of the population happened seven years later, after a sudden increase in the military offensive in 2008. I use the lag period between the beginning of the conflict (September 2001) and the onset of the IDP crisis (2008) to isolate the potential impact of the conflict. There is no effect on the number of polio cases before 2008 (see panel A of Table A.12).[15]

The peculiarity of this paper's setting is that most IDPs moved to districts within historical *Pashtunistan*. Suppose the large IDP inflows create the main effects. In that case, we should not observe an effect when comparing my counterfactual (red polygons in Figure A.17) to other non-Pashtu districts non-included in my baseline sample (white polygons in Figure A.17). The results of panel B of Table A.13 align with this assumption.

[15]The results do not change when I use November 2004 as my falsification treatment timing.

7.2 Treatment and Counterfactual Definition

The definition of my treatment relies on the historical border of *Pashtunistan*. As highlighted in section 4.1, I remove from my baseline sample the districts where only a share of their territory falls within *Pashtunistan*. These districts correspond to the red-dashed polygons in Figure A.17. I add these districts to my treatment definition. I show in Panel A of Table A.14 that although the magnitude of the effects decreases, the points estimates are statistically significant at the five per cent level. To facilitate a quick return and due to cultural barriers, IDP families settled in districts near F.A.T.A. Still, a minor share of IDPs moved to further districts within the historical *Pashtunistan*. The results of Table A.15 support this fact by presenting a very small or no effect when restricting the treatment to districts overlapping the Pashtu line.

I conduct two similar exercises to validate my counterfactual definition. First, I use the overlapping district to the Pashtu line as an alternative control group. The findings of Table A.15 would imply that we should observe an adverse and significant effect when using this alternative counterfactual. Table A.16 test and validate this hypothesis. Using the alternative counterfactual allows me to control for potential unobservables related to the Pashtu traditions and culture. Second, I show that the results hold when using the non-Pashtu districts not included in my baseline counterfactual as an alternative control group. See Table A.17 for further details.

7.3 Alternative Specification and Potential Cofounders

Polio cases have changed in the country over time, where the health response is the provincial government's responsibility. Thus, in my baseline specification (equation 1), I control for a year-month fixed effect to account for seasonal shocks standard across all districts in Pakistan and a provincial fixed effect to control for time-invariant characteristics within a district. However, my results are robust to alternative specifications. When controlling for the division fixed effect, only year fixed effect, or only province fixed effect, the magnitude of the effect is precisely the same. It only changes the significance level at the ten per cent level and increases the standard errors when I control for district-fixed effects (see panels A, B and C of Table A.18). Panel D shows how the standard errors decrease when I do not cluster them.

Finally, it could be a concern that IDP families chose their host community based on previous or existing polio cases number. Hence, a potential reverse causality could threaten my identification. Table A.19 rules out this hypothesis using the aggregate district polio cases from 2001 to 2007.

8 Conclusion

The increasing number of internally displaced people (IDP) poses new challenges to eradicating polio globally. This paper provides evidence that communities that received the IDP population increased their polio incidence, where the intensity of the inflow matters substantially.

Three transmitting channels can explain these findings: First, although a vaccination programme targeting children on the move reduced the gaps in the immunization rates between IDP and native children, the vaccination rates at the national level are too low to stop the spread of the virus when new polio cases pop up in host communities. Second, IDP families migrate to the poorest and most marginalized communities where other migrants are or can afford to live. The precarious health condition in the new destinations facilitates the spread of the virus. Third, the inflow of the IDP population could have congested health services in host communities.

The situation in Pakistan is not unique. In 2022, polio cases emerged in Malawi and Mozambique– two countries free of the virus for decades. Both countries are the scenario of a forced population movement from the conflicts ongoing in Northern Mozambique and the Eastern Democratic Republic of Congo, respectively. Protected children from diseases are far more likely to have the opportunity to thrive, the chance to learn and the ability to live healthy lives (UNICEF 2023). Therefore, three critical policy implications emerge from this paper. First, millions of forcibly displaced children migrate to camps or host communities. Since families in these settings are often transient, monitoring vaccination rates among these communities make it much harder to reach children with the necessary vaccines. Reaching the hard-to-reach- such as children from mobile and forced migrant populations or in conflict zones- should be a public priority (CDC 2021; UNICEF 2023). Second, poor communities are the host communities of most of the IDPs. An effort to better integrate the IDP population into the health services and labour market should be made to improve the conditions in which they live. Finally, the inflow of new population comes with increased demand for health services. Even if the increase in the demand is modest, in locations where the health delivery or capacity is weak, it can congest the local health services. It is essential to reinforce host communities' health workforce and infrastructure, so locals and newcomers can access health services equally.

References

Abbas, Hassan (2008). "A Profile of Tehrik-I-Taliban Pakistan". In: *CTC Sentinel. West Point, NY: Combating Terrorism Center. 1 (2): 1–4. Archived from the original on 1 January 2017. Retrieved 8 November 2008.*

Adda, Jérôme (2016). "Economic Activity and the Spread of Viral Diseases: Evidence from High Frequency Data". In: *The Quarterly Journal of Economics, Volume 131, Issue 2, May 2016, Pages 891–941.*

Baez, Javier E. (2011). "Civil wars beyond their borders: The human capital and health consequences of hosting refugees". In: *Journal of Development Economics 96 (2011) 391–408.*

Becker, Sascha O. and Andreas Ferrara (2019). "Consequences of forced migration: A survey of recent findings". In: *Labour Economics. Volume 59, August 2019, Pages 1-16.*

Bezhan, Faridullah (2014). "The Pashtunistan Issue and Politics in Afghanistan, 1947-1952". In: *Middle East Journal, Vol. 68, No. 2 (Spring 2014), pp. 197-209 (13 pages).*

Blattman, Christipher and Edward Miguel (2010). "Civil War". In: *Journal of Economic Literature 2010, 48:1, 3–57.*

Bundervoet, Tom, Philip Verwimp, and Richard Akresh (2009). "Health and Civil War in Rural Burundi". In: *Journal of Human Resources, March 2009, 44 (2) 536-563.*

CDC (2021). "Polio Vaccine: What you need to know". In: *Centers for Disease Control and Prevention, 2021.*

Devkota, Bhimsen and Edwin R van Teijlingen (2010). "Understanding effects of armed conflict on health outcomes: the case of Nepal". In: *Conflict and Health volume 4, Article number: 20 (2010).*

Din, Najam U (2010). "Internal Displacement in Pakistan: Contemporary Challenges". In: *Human Rights Commission of Pakistan. Printed at Qasim Press Railway Road, Lahore.*

GTD (2021). "Global Terrorism Database". In: *Access at https://www.start.umd.edu/gtd/.*

Hussain, Shoaib F. et al. (2016). "Eradicating polio in Pakistan: an analysis of the challenges and solutions to this security and health issue". In: *Global Health. 2016 Oct 12;12(1):63.*

Ibáñez, Ana María, Andrés Moya, et al. (2023). "Least Protected, Most Affected: Impacts of Migration Regularization Programs on Pandemic Resilience". In: *AEA Papers and Proceedings. 2023. Forthcoming.*

Ibáñez, Ana María, Sandra V. Rozo, and María J. Urbina (2021). "Forced Migration and the Spread of Infectious Diseases". In: *Journal of Health Economics. Volume 79, September 2021, 102491.*

IDMC (2015). "Solutions to displacement elusive for both new and protracted IDPs". In: *Geneva: Internal Displacement Monitoring Centre.*

Malik, Adeel, Rinchan Ali Mirza, and Faiz Ur Rehman (2023). "Frontier Governmentality". In: *Working Paper.*

Martinez-Bravo, Monica and Andreas Stegmann (2022). "In Vaccines We Trust? The Effects of the CIA's Vaccine Ruse on Immunization in Pakistan". In: *Journal of the European Economic Association, Volume 20, Issue 1, February 2022, Pages 150–186.*

Montalvo, Jose G. and Marta Reynal-Querol (2007). "Fighting against Malaria: Prevent wars while waiting for the "Miraculous" Vaccine". In: *The Review of Economics and Statistics, February 2007, 89(1): 165–177.*

Mushtaq, Asim et al. (2015). "Polio in Pakistan: Social constraints and travel implications". In: *Travel Medicine and Infectious Disease (2015) 13, 360e366.*

New-America (2021). "World of Drones". In: *Access at https://www.newamerica.org/.*

Oster, Emily (2012). "Routes of Infection: Exports and Hiv Incidence in Sub-Saharan Africa". In: *Journal of the European Economic Association, Volume 10, Issue 5, 1 October 2012, Pages 1025–1058.*

Pérez-Sindín, Xaquín S., Tzu-Hsin Karen Chen, and Alexander V. Prishchepov (2021). "Are night-time lights a good proxy of economic activity in rural areas in middle and low-income countries? Examining the empirical evidence from Colombia". In: *Remote Sensing Applications: Society and Environment. Volume 24, November 2021, 100647.*

Phadera, Lokendra (2021). "Unfortunate Moms and Unfortunate Children: Impact of the Nepali Civil War on Women's Stature and Intergenerational Health". In: *Journal of Health Economics. Volume 76, March 2021, 102410.*

Rahim, Shabina, Zubair Ahmad, and Jamshid Abdul-Ghafar (2022). "The polio vaccination story of Pakistan". In: *Vaccine. Volume 40, Issue 3, 24 January 2022, Pages 397-402.*

UNDESA (2015). "World Population Prospects: The 2015 Revision". In: *United Nations Department of Economic and Social Affairs.*

UNHCR (2017). "Obtaining representative data on IDPs: challenges and recommendations". In: *UNHCR Statistics Technical Series: 2017/1.*

— (2022). "United Nations High Commissioner for the Refugees Data". In: *Access at https://www.unhcr.org/.*

UNICEF (2019). “National Emergency Action Plan 2018-2019”. In: *UNICEF*.

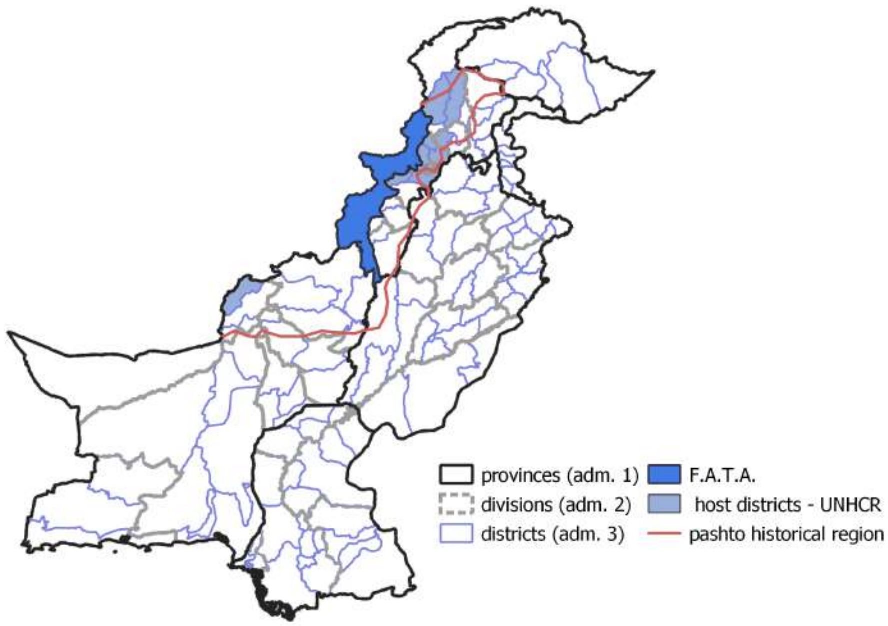
— (2023). “The State of the World’s Children 2023: For every child, vaccination”. In: *UNICEF Innocenti – Global Office of Research and Foresight, Florence, April 2023*.

Vogt, Manuel et al. (2015). “Integrating Data on Ethnicity, Geography, and Conflict: The Ethnic Power Relations Data Set Family”. In: *Journal of Conflict Resolution 59(7): 1327–42*.

WHO (2022). “Polio vaccines: WHO position paper – June 2022”. In: *World Health Organization, 2022, 97, 277–300*.

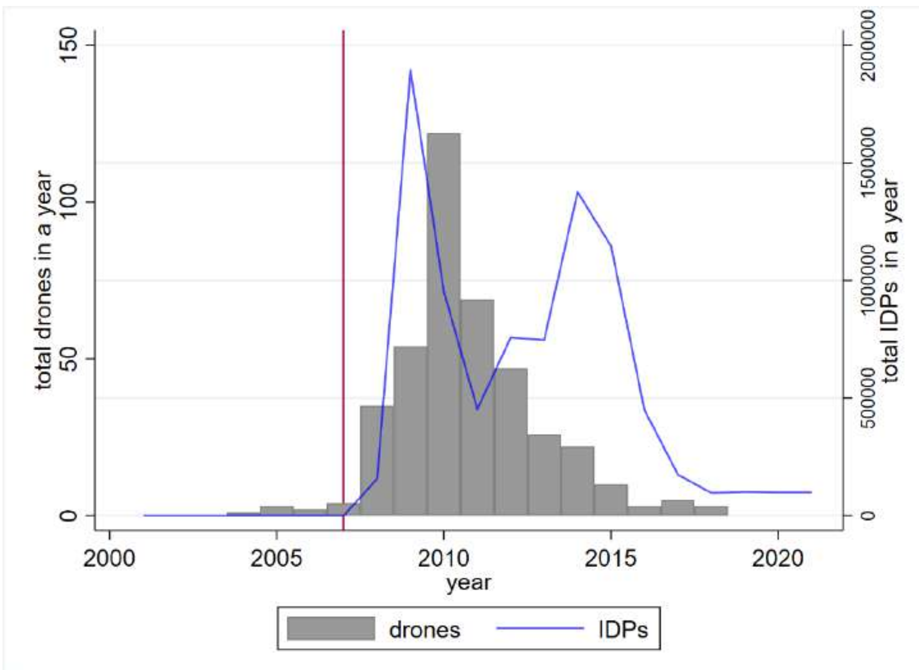
Figures and Tables

Figure 1: Host destinations and Pashtunistan



Note: This figure shows the spatial distribution of the main host districts in Pakistan. In light blue, I show the districts which received the internally displaced population from the Federally Administered Tribal Areas (F.A.T.A.) recorded by UNHCR (UNHCR 2022). The region of F.A.T.A. is in dark blue. The red line illustrates the pre-colonial region of *Pashtunistan.* Black polygons correspond to the provinces (the first administrative division in Pakistan). Grey polygons correspond to division (the second administrative division). And the white polygons with purple lines are the districts (the third administrative division).

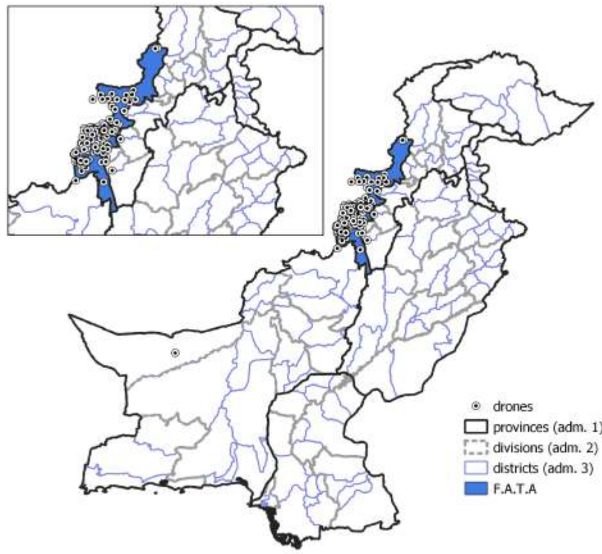
Figure 2: Total drones strikes and IDP population (2000-2022)



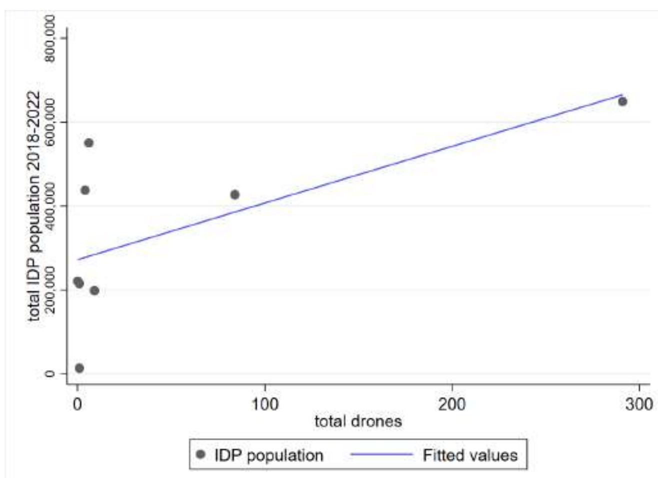
Note: This figure shows the yearly number of drones and internally displaced persons (IDP) from 2001 to 2022. The grey bars show the number of drones and the blue line to the number of IDPs. The vertical red line corresponds to 2007.

Figure 3: Drones as a migration push factor

a) Drone strikes locations

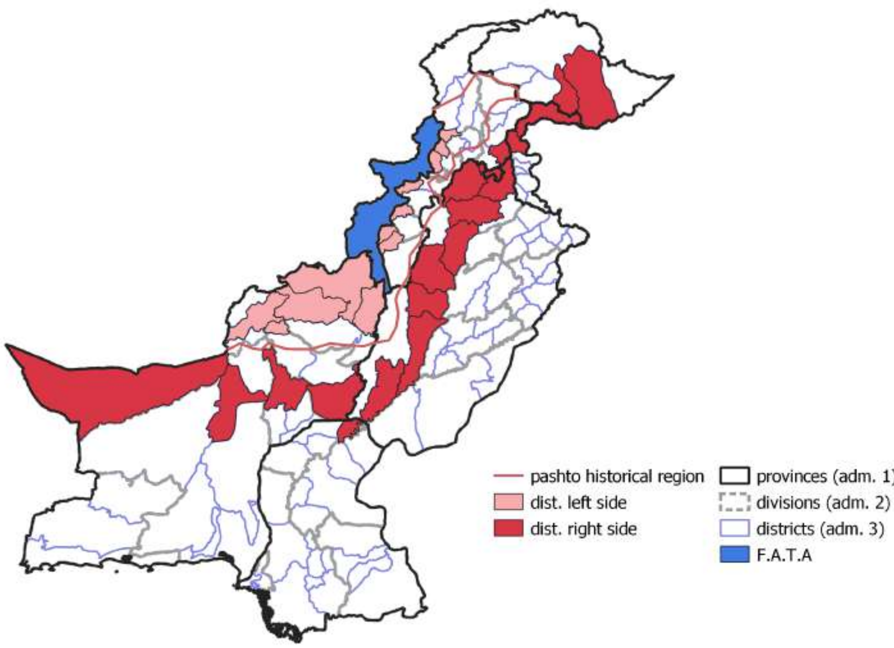


b) Correlation between the number of drones and IDPs



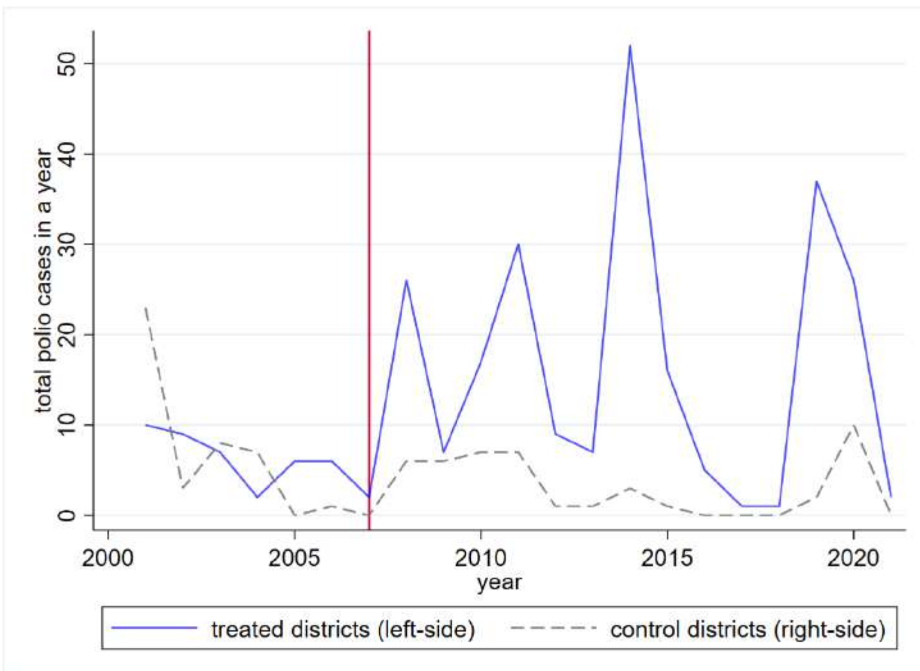
Note: This figure shows the relationship between the intensity of drone strikes and migration. Figure A illustrates the spatial distribution of drones from 2001 to 2022 in Pakistan. The blue polygons correspond to the Federally Administered Tribal Areas (F.A.T.A.) region. Figure B plots the correlation between the district's number of drones and the aggregate IDPs from 2001 to 2022.

Figure 4: Treated and control districts



Note: This figure shows treated (host) and control (non-host) districts. To define them, I use the spatial distribution of districts relative to the pre-colonial region of *Pashtunistan*. The red line corresponds to the *Pashtunistan*'s border. Districts whose territory falls within the pre-colonial region of *Pashtunistan* are host districts. Non-host districts are those whose territory is outside *Pashtunistan* but adjacent to the historical border.

Figure 5: Polio cases (2001-2022)



Note: This figure plots the yearly polio cases in treated and control districts. Treated districts are the host districts, and control districts are the non-host districts. Districts whose territory falls within the pre-colonial region of *Pashtunistan* are host districts. Non-host districts are those whose territory is outside *Pashtunistan* but adjacent to the historical border.

Table 1: Effect of IDP population on polio

	(1)	(2)	(3)	(4)	(5)
			PANEL A: At least one polio case		
VARIABLES	polio	polio	polio	polio	polio
2007 x Host district	0.038***	0.038**	0.053**	0.056*	0.067**
	(0.009)	(0.012)	(0.018)	(0.023)	(0.022)
Observations	8,184	8,184	8,184	6,600	5,040
Number of provinces		7	7	4	7
			PANEL B: Number of polio cases		
VARIABLES	polio cases	polio cases	polio cases	polio cases	polio cases
2007 x Host district	0.081***	0.081**	0.105**	0.114	0.117**
	(0.015)	(0.032)	(0.042)	(0.054)	(0.045)
Observations	8,184	8,184	8,184	6,600	5,040
Number of provinces		7	7	4	7
			PANEL C: Number of polio cases per 100,000 inhabitants		
VARIABLES	polio pop.den.	polio pop.den.	polio pop.den.	polio pop.den.	polio pop.den.
2007 x Host district	0.006575***	0.006575*	0.006793**	0.007175	0.009553***
	(0.001872)	(0.002769)	(0.002445)	(0.003326)	(0.001200)
Observations	7,128	7,128	7,128	6,600	4,740
Number of provinces		5	5	4	5
Province FE	No	Yes	Yes	Yes	Yes
Time FE	No	Yes	Yes	Yes	Yes
Controls	No	No	Yes	Yes	Yes

Note: This Table presents the impacts of the IDP inflows on district polio prevalence in host districts compared to non-host district. I use the spatial distribution of districts with respect the pre-colonial region of *Pashtunistan* to define host and non-host districts. Districts whose territory fall within the pre-colonial region of *Pashtunistan* are defined as host districts. Non-host districts are the district whose territory is outside *Pashtunistan*, but are adjacent to the historical border. The treatment timing starts from 2008. Observations are at the district and month level from 2001 to 2022. The baseline specification is presented in equation (1). Column (1) presents the results without province, year-month fixed effects and covariates. Column (2) includes province and year-month fixed effects. Column (3) controls for nightlight intensity and total vaccination campaigns. Columns (4) controls for pre-treatment district-covariates (the average number of children under five, the average number of members in a household, and the total share of the literate population from 1973, 1981 and 1998 Population Census). Column (5) control instead for contemporary characteristics (the average number of children under five, the average number of members in a household, shared households with piped water, and shared households with a finished floor). This Table present three different outcomes: at least one case of polio (panel A), total number of polio cases (panel B) and polio cases per 100,000 inhabitants (panel C). Standard errors are clustered at the province level. Robust standard errors are in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 2: Effect of predicted IDP inflow on polio

	(1)	(2)	(3)	(4)	(5)
	PANEL A: At least one polio case				
VARIABLES	polio	polio	polio	polio	polio
Predicted Inflow	0.022901***	0.021437***	0.019206***	0.019725***	0.018128***
	(0.003250)	(0.003127)	(0.001494)	(0.001193)	(0.002035)
Observations	8,184	8,184	8,184	6,600	5,040
Number of provinces		7	7	4	7
	PANEL B: Number of polio cases				
VARIABLES	polio cases	polio cases	polio cases	polio cases	polio cases
Predicted Inflow	0.041253***	0.041242***	0.037685***	0.040281***	0.033821***
	(0.007991)	(0.005771)	(0.003034)	(0.002563)	(0.003542)
Observations	8,184	8,184	8,184	6,600	5,040
Number of provinces		7	7	4	7
	PANEL C: Number of polio cases per 100,000 inhabitants				
VARIABLES	polio pop.den.	polio pop.den.	polio pop.den.	polio pop.den.	polio pop.den.
Predicted Inflow	0.001876***	0.001427***	0.001379***	0.001432**	0.001396**
	(0.000470)	(0.000161)	(0.000208)	(0.000418)	(0.000485)
Observations	7,128	7,128	7,128	6,600	4,740
Number of provinces		5	5	4	5
Province FE	No	Yes	Yes	Yes	Yes
Time FE	No	Yes	Yes	Yes	Yes
Controls	No	No	Yes	Yes	Yes

Note: This Table presents the results on polio prevalence with a continuous treatment: the predicted IDP inflows. I construct my predicted inflows measure as the interaction of the inverse distance of each district to the closest F.A.T.A border (district variation) and the total yearly number of IDP population (annual variation). The treatment timing starts from 2008. Observations are at the district and month level from 2001 to 2022. The baseline specification is presented in equation (1). Column (1) presents the results without province, year-month fixed effects and covariates. I restrict the sample to the districts defined as host or non-host district. Column (2) includes province and year-month fixed effects. Column (3) controls for nightlight intensity and total vaccination campaigns. Columns (4) controls for pre-treatment district-covariates (the average number of children under five, the average number of members in a household, and the total share of the literate population from 1973, 1981 and 1998 Population Census). Column (5) control instead for contemporary characteristics (the average number of children under five, the average number of members in a household, shared households with piped water, and shared households with a finished floor). This Table present three different outcomes: at least one case of polio (panel A), total number of polio cases (panel B) and polio cases per 100,000 inhabitants (panel C). Standard errors are clustered at the province level. Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 3: Effect of IDP population on vaccination against polio

	(1)	(2)	(3)	(4)	(5)
VARIABLES	vaccinated	vaccinated	vaccinated	vaccinated	vaccinated
	PANEL A: Cohort variation				
2007 x Host district	0.004	-0.064*	-0.080*	-0.067**	-0.083*
	(0.011)	(0.033)	(0.041)	(0.032)	(0.041)
	PANEL B: Heterogenity between IDP and native children				
2007 x Host district	0.004	-0.064*	-0.080*	-0.067**	-0.082*
	(0.011)	(0.033)	(0.041)	(0.031)	(0.041)
2007 x Host district x IDP	-0.004	-0.041	-0.041	-0.046	-0.046
	(0.075)	(0.109)	(0.109)	(0.109)	(0.109)
Observations	10,608	10,608	10,608	10,563	10,563
Province FE	No	Yes	Yes	Yes	Yes
Cohort FE	No	Yes	Yes	Yes	Yes
Controls	No	No	Yes	Yes	Yes
Number of districts		31	31	31	31
	PANEL C: Districts variation				
Predicted Inflow	0.021***	-0.004	-0.005	-0.012	-0.011
	(0.005)	(0.015)	(0.016)	(0.017)	(0.017)
Observations	10,608	10,608	10,608	10,563	10,563
Province FE	No	Yes	Yes	Yes	Yes
Cohort FE	No	Yes	Yes	Yes	Yes
Controls	No	No	Yes	Yes	Yes
Number of provinces		7	7	7	7

Note: This Table presents the impacts of the IDP inflows on vaccination behaviours at individual level. The dependent variable is a binary variable for being vaccinated, coded to one if the children is vaccinated. I use individual-level data on vaccination against polio and the date of birth from the Demographic and Health Survey (DHS) from 1998 to 2017. Children born from January 2008 are exposed to the treatment. The baseline specification is presented in equation (2), where I rely on within-district cohort variation. Column (1) presents the results without district, cohort fixed effects and covariates. Column (2) includes district and cohort fixed effects. Column (3) controls for nightlight intensity and total vaccination campaigns. Columns (4) controls for the head of household, urban location, and gender of the child. Column (5) control for the full set of covariates included in columns (3) and (4). Standard errors are clustered at the province level. Panel A present the baseline results. Panel B shows the estimates adding an interaction if a child is IDP. Panel C control for province fixed effects to compare within-cohort children between host and non-host districts. I use the spatial distribution of districts with respect the pre-colonial region of *Pashtunistan* to define host and non-host districts. Districts whose territory fall within the pre-colonial region of *Pashtunistan* are defined as host districts. Non-host districts are the district whose territory is outside *Pashtunistan*, but are adjacent to the historical border. Robust standard errors are in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 4: Effect of IDP population on household health conditions

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	water piped	toilet	floor	total children	total members	working
	PANEL A: Cohort variation					
2007 x Host district	-0.221***	-0.148	-0.089	-0.271	-0.461	-0.004
	(0.072)	(0.091)	(0.075)	(0.248)	(0.886)	(0.039)
	PANEL B: Cohort variation - predicted inflow					
Predicted Inflow	-0.093***	0.003	0.005	-0.051	0.015	-0.009
	(0.017)	(0.023)	(0.031)	(0.056)	(0.205)	(0.015)
	PANEL C: Heterogenity between IDP and native children					
Predicted Inflow	-0.095***	0.004	0.005	-0.047	0.002	-0.008
	(0.017)	(0.023)	(0.031)	(0.057)	(0.208)	(0.015)
Predicted Inflow x IDP	0.058**	-0.031	-0.022	-0.130*	0.434*	-0.023***
	(0.022)	(0.030)	(0.032)	(0.072)	(0.229)	(0.006)
Observations	10,623	10,623	7,865	10,623	10,623	10,578
Number of districts	31	31	31	31	31	31

Note: This Table presents the impacts of the IDP inflows on household characteristics. There are six different dependent variables: access to drinkable water (column (1)), access to a toilet (column (2)), floor quality (column (3)), number of children under five (column (4)), households member (column (5)), and head of the household working (column (6)). The dependent variables are a binary, coded to one if the household has a certain characteristic. I use individual-level data on household characteristics and the date of the interview from the Demographic and Health Survey (DHS) from 1998 to 2017. Households interviewed from January 2008 are exposed to the treatment. The baseline specification is presented in equation (3), where I rely on within-district household variation. Panel A present the baseline results. Panel B shows the estimates with a continuous treatment: the predicted IDP inflows. Panel C adds an interaction if a child is IDP. Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 5: Effect of IDP inflow on the demand and supply of health services

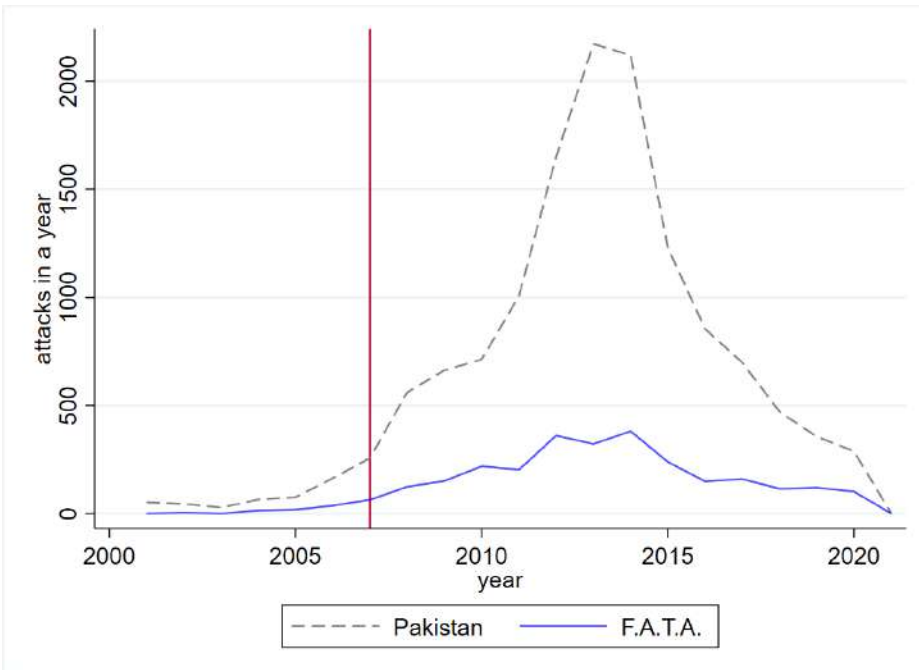
	(1)	(2)	(3)	(4)	(5)	(6)
			PANEL A: Demand health services			
VARIABLES	doctor prenatal	d. prenatal	d. prenatal	doctor postnatal	d. postnatal	d. postnatal
2007 x Host district	-0.037			0.032		
	(0.056)			(0.057)		
Predicted Inflow		0.003	0.003		0.014	0.013
		(0.014)	(0.014)		(0.017)	(0.017)
Predicted Inflow X IDP			0.007			0.019**
			(0.008)			(0.006)
Observations	10,623	10,623	10,623	10,623	10,623	10,623
District FE	Yes	No	No	Yes	No	No
Province FE	No	Yes	Yes	No	Yes	Yes
Controls	No	No	No	No	No	No
Number of districts	31	31	31	31	31	31
Number of provinces	7	7	7	7	7	7
			PANEL B: Number of polio cases			
VARIABLES	polio act.	polio act.	polio act.	polio act.	polio act.	
Predicted Inflow	0.112559***	0.066885***	0.060775***	0.065066***	0.062826***	
	(0.011082)	(0.004376)	(0.004399)	(0.009431)	(0.009756)	
Observations	8,184	8,184	8,184	6,600	5,040	
Province FE	No	Yes	Yes	Yes	Yes	
Time FE	No	Yes	Yes	Yes	Yes	
Controls	No	No	Yes	Yes	Yes	

Note: This Table presents the impacts of the IDP inflows on the demand and supply of health services. In Panel A I estimate equation (2) on individual-level data on prenatal (columns (1), (2), and (3)) and postnatal doctor assistance (columns (4), (5), and (6)). I exploit the within district cohort variation to identify the effects. I include province-fixed effects in columns (2), (3), (5), and (6). Panel B shows the estimates on the predicted IDP inflows on district-level data on polio vaccination campaigns from the Polio Eradication Program. The baseline specification is presented in equation (1) Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Appendices

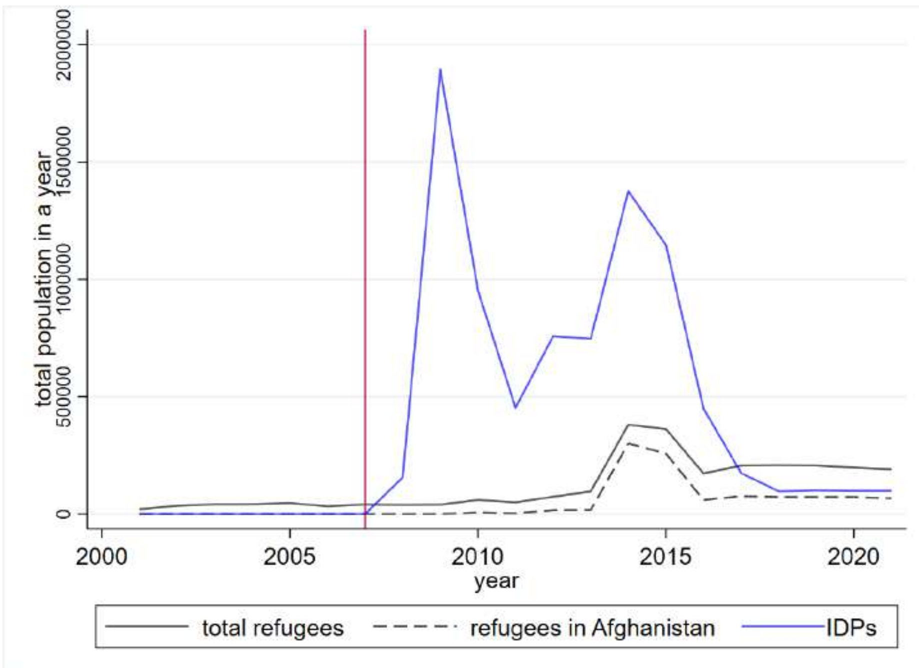
A Additional Figures

Figure A.1: Terrorist attacks (2001-2022)



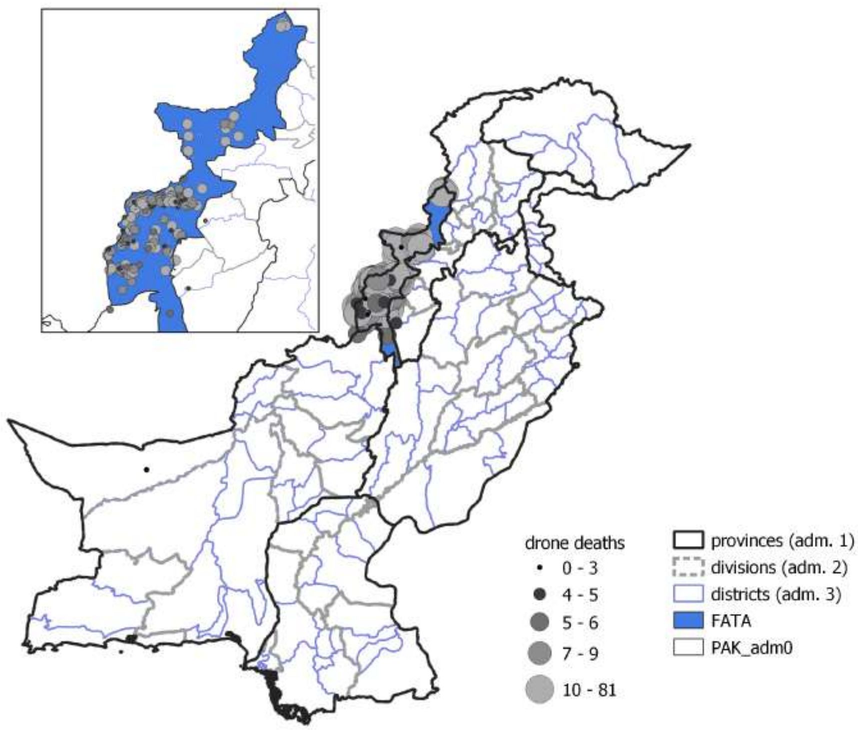
Note: This figure shows the yearly number of terrorist attacks from 2001 to 2022. The grey dashed line for Pakistan and the blue line for F.A.T.A. The vertical red line corresponds to 2007. Source. The Global Terrorism Database - G.T.D. (GTD 2021).

Figure A.2: Forcibly displaced population within and outside Pakistan (2001-2022)



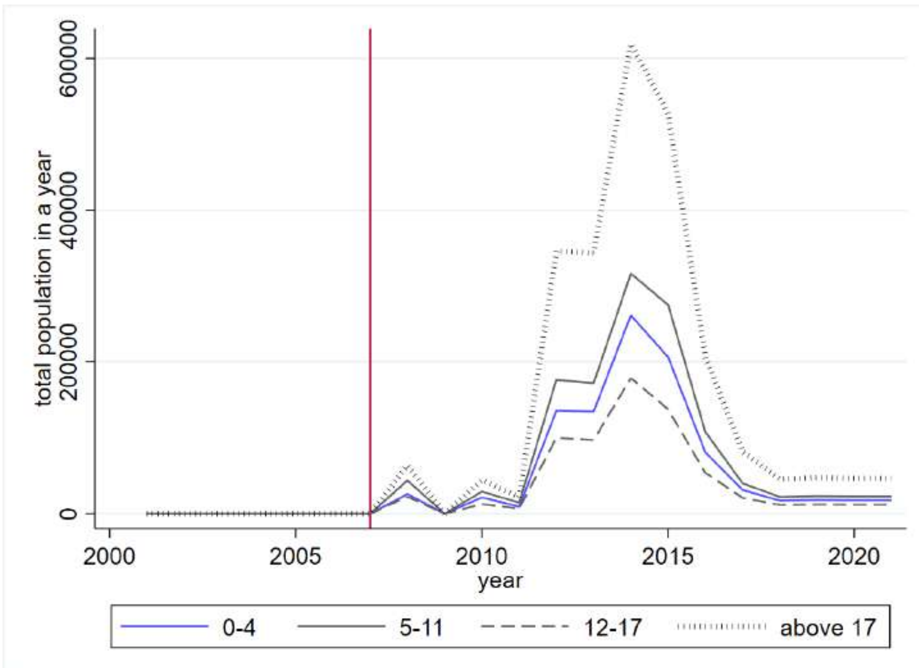
Note: This figure shows the yearly displaced population from Pakistan from 2001 to 2022. The blue line corresponds to the internally displaced persons (IDP). The black line shows the number of Pakistani refugees worldwide. And the black dashed line of the Pakistani refugees in Afghanistan. The vertical red line corresponds to 2007. Source. The United Nations High Commissioner for Refugees - U.N.H.C.R. (UNHCR 2022).

Figure A.3: Total deaths by drones (2001-2022)



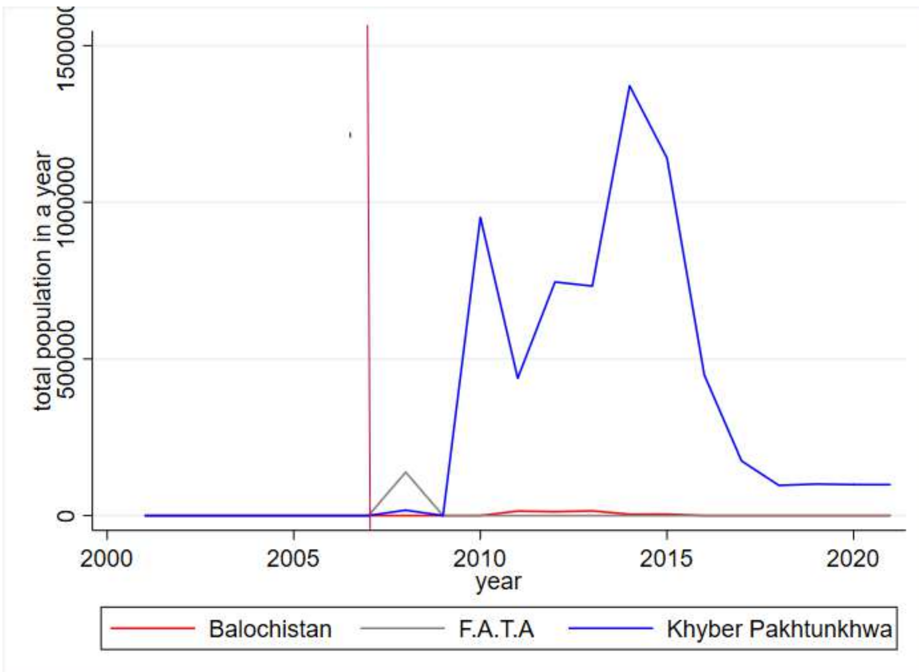
Note: This figure shows the spatial distribution of deaths associated with each drone strike from 2001 to 2022. The higher the dot higher is the total number of deaths. Source. The World of Drones Database developed by New America (New-America 2021).

Figure A.4: Internally Displaced Persons by age (2001-2022)



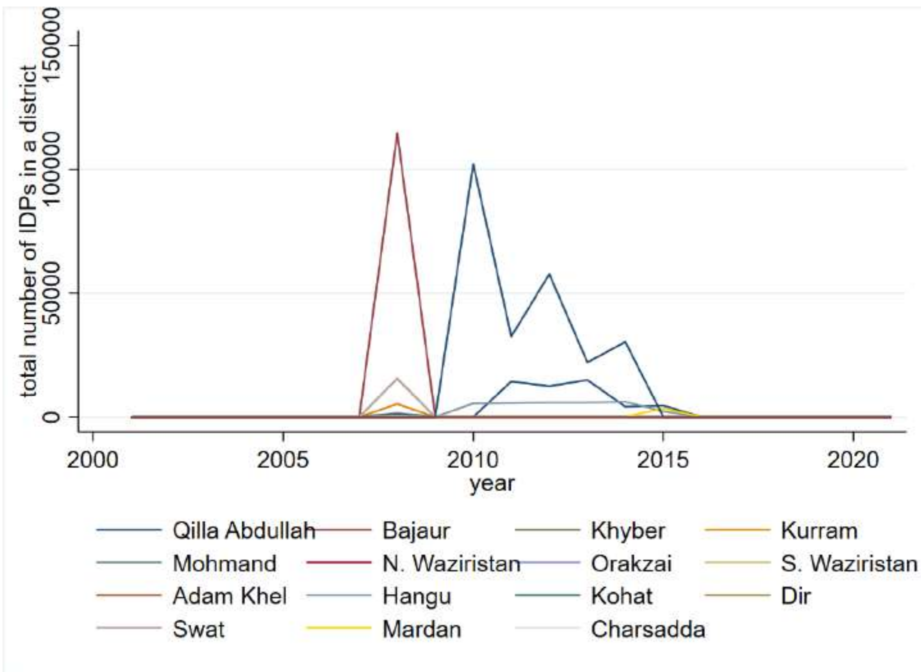
Note: This figure shows the yearly internally displaced population by age from 2001 to 2022. The blue line corresponds to the ages 0-4, the black line to the ages 5-11, the black dashed line to the ages 12-17 and the black pointed line to the ages above 17. The vertical red line corresponds to 2007. Source. The United Nations High Commissioner for Refugees - U.N.H.C.R. (UNHCR 2022).

Figure A.5: Internally Displaced Population by province



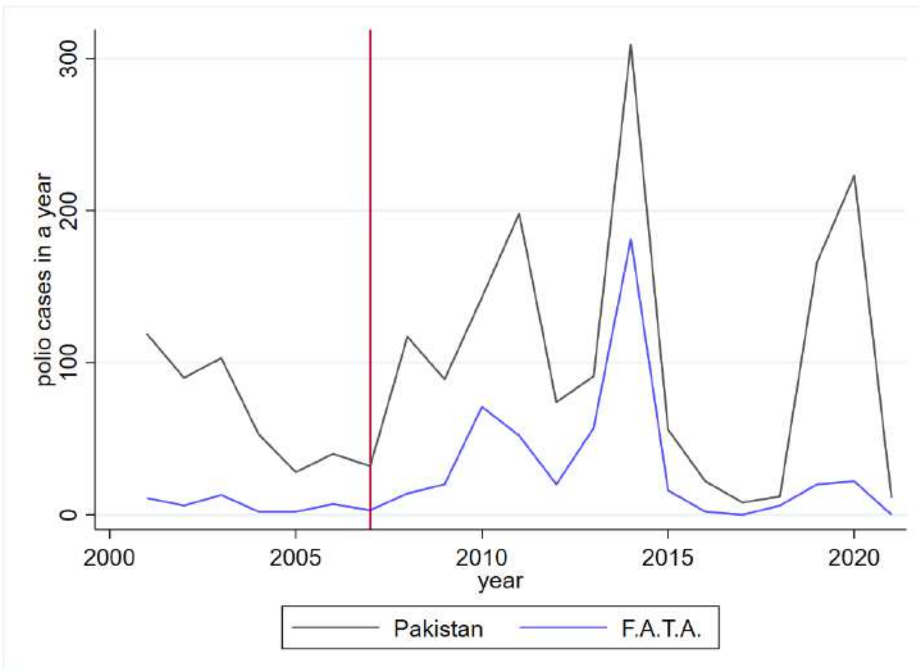
Note: This figure shows the yearly internally displaced population by province from 2001 to 2022. The red line corresponds to the province of Balochistan, border to Southern F.A.T.A.. The grey line are the IDPs in F.A.T.A. The blue line corresponds to Khyber Pakhtunkhwa province. Khyber Pakhtunkhwa is the border of the Eastern and Northern F.A.T.A. The vertical red line corresponds to 2007. Source. The United Nations High Commissioner for Refugees - U.N.H.C.R. (UNHCR 2022).

Figure A.6: Internally Displaced Population by destination district (2001-2022)



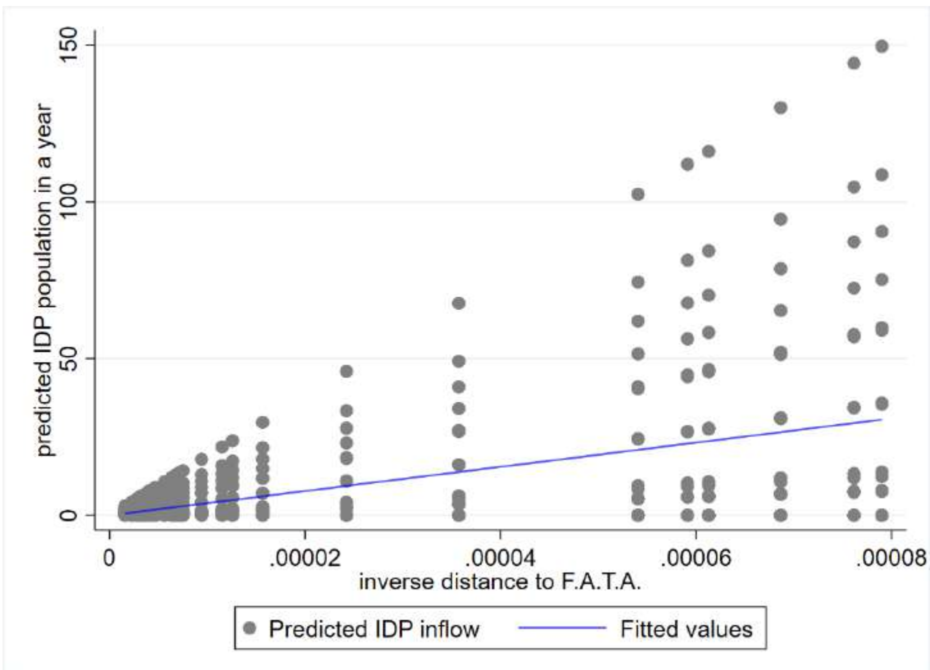
Note: This figure shows the yearly internally displaced population by district from 2001 to 2022. Source. The United Nations High Commissioner for Refugees - U.N.H.C.R. (UNHCR 2022).

Figure A.7: Polio cases in Pakistan and F.A.T.A. (2001-2022)



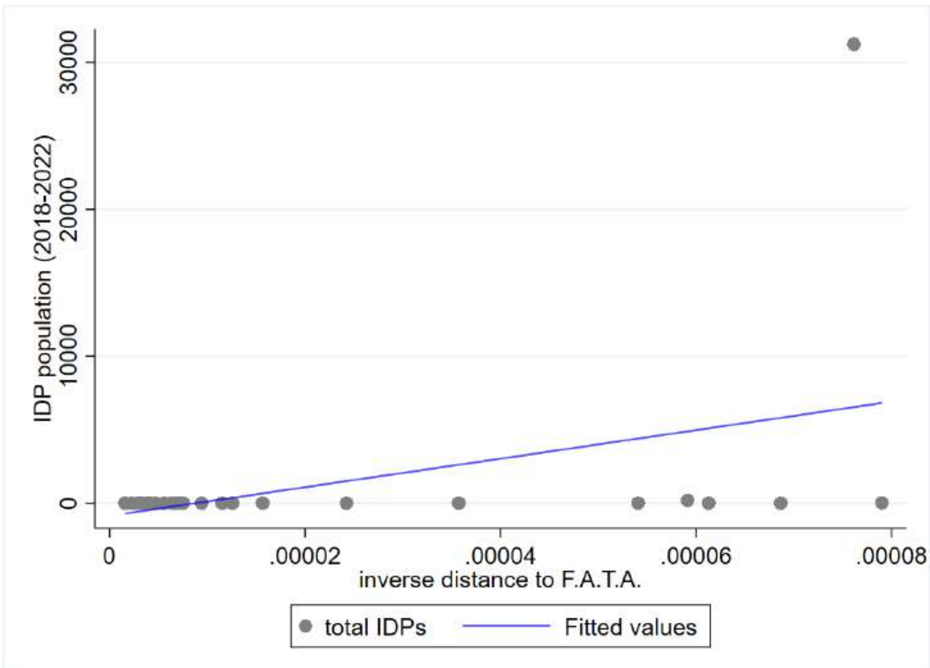
Note: This figure plots the yearly polio cases in Pakistan and F.A.T.A. Treated districts are the host districts, and control districts are the non-host districts. Source. The Polio Eradication Program established by the World Health Organization (WHO).

Figure A.8: Inverse distance and predicted inflow



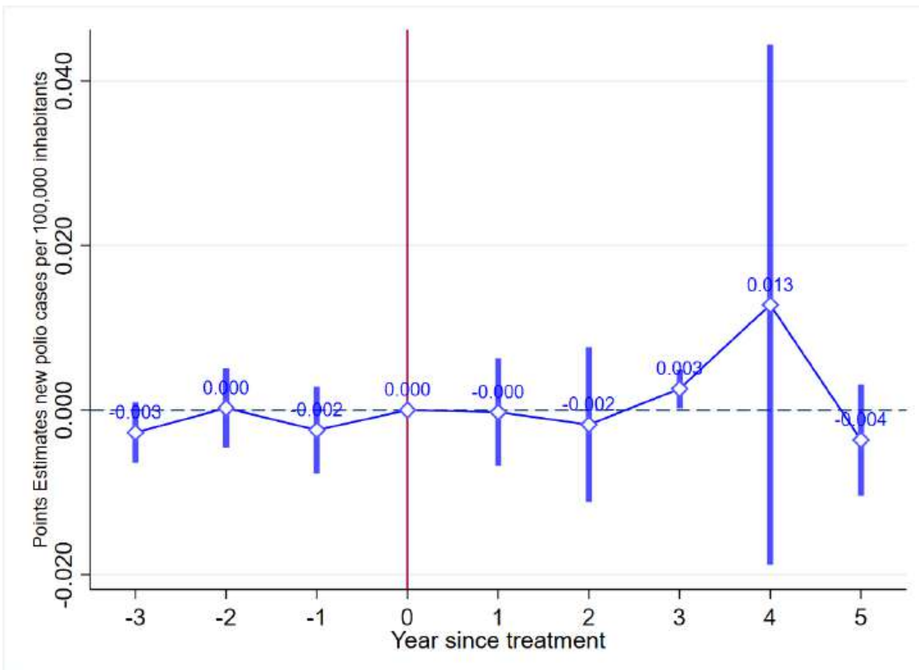
Note: This figure shows the correlation between the predicted inflow measure and the inverse distance to the closest F.A.T.A. border. The predicted inflow measure is equal to the interaction of the inverse distance of each district to the nearest F.A.T.A. border (district variation) and the total yearly number of IDP population (annual variation).

Figure A.9: Inversed distance and reported IDPs



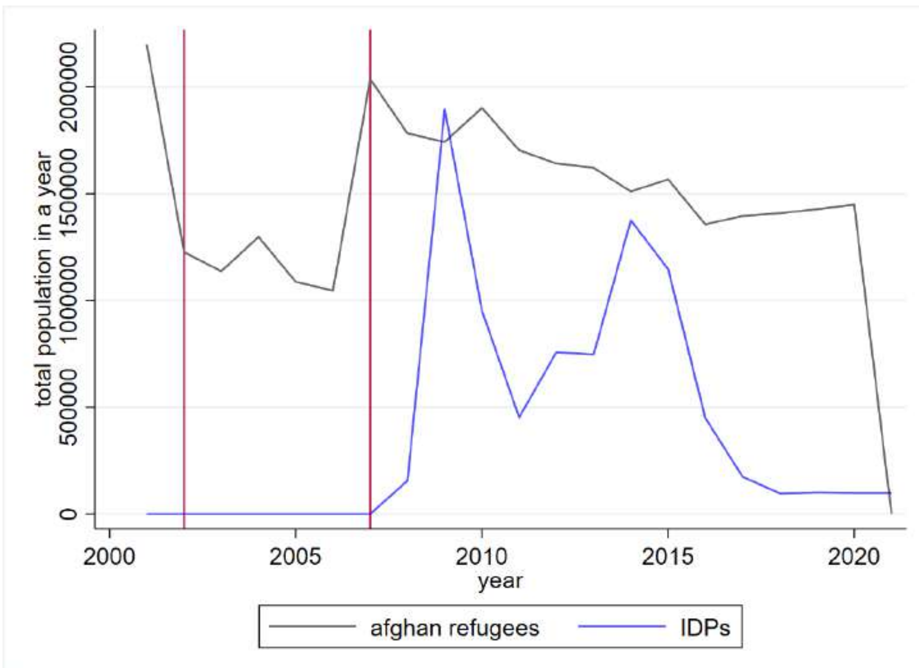
Note: This figure shows the correlation between the actual IDP inflow and the inverse distance to the closest F.A.T.A. border. The IDP information comes from the United Nations High Commissioner for Refugees - U.N.H.C.R. (UNHCR 2022).

Figure A.10: Event study



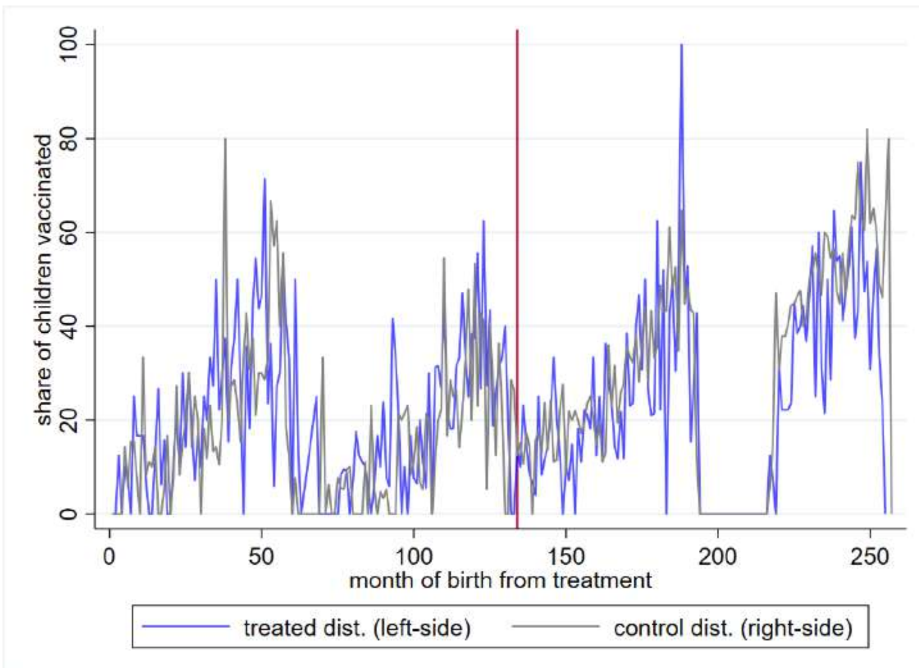
Note: Figure A.10 plots the event and year coefficient from estimating equation 1 using the new polio cases per inhabitant as the dependent variable. The confidence intervals are 95%. Polio outcomes come from the Polio Eradication Program established by the World Health Organization (WHO). The omitted category is T-0, the year 2007. The dataset is in a year-district panel format. Treatment is defined at the year level.

Figure A.11: Afghan refugees (2001-2022)



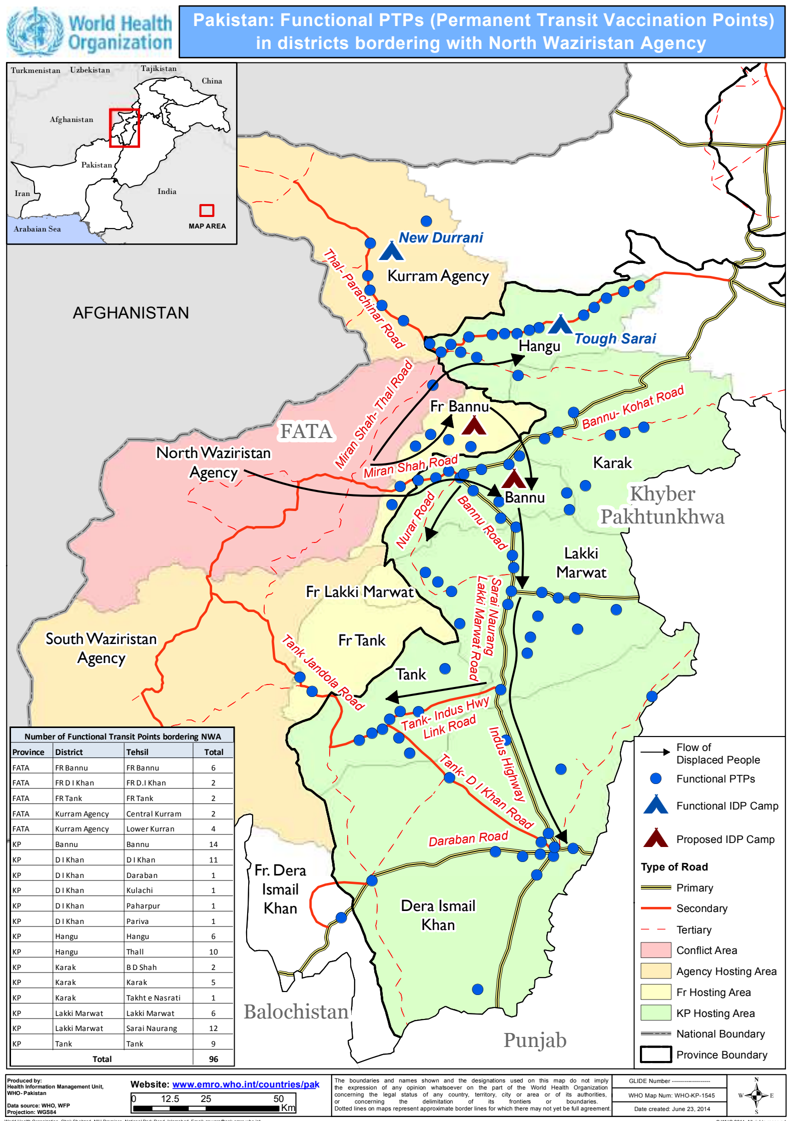
Note: This figure shows the yearly internally displaced population and Afghan refugees in Pakistan from 2001 to 2022. The blue line corresponds to the internally displaced persons (IDP). The black line indicates the number of Afghan refugees in Pakistan. The vertical red line corresponds to 2007. Source. The United Nations High Commissioner for Refugees - U.N.H.C.R. (UNHCR 2022).

Figure A.12: Inmunization rates (2001-2022), host vs non-host



Note: This figure plots the share of children vaccinated in treated and control districts by the birth cohort. Treated districts are the host districts and control districts are the non-host districts. Districts whose territory falls within the pre-colonial region of *Pashtunistan* are host districts. Non-host districts are those whose part is outside *Pashtunistan* but adjacent to the historical border. The vertical red line corresponds to December 2007. Source. Demographic and Health Survey (DHS).

Figure A.13: Permanent Transit Points (PTPs) to vaccinate children on the move

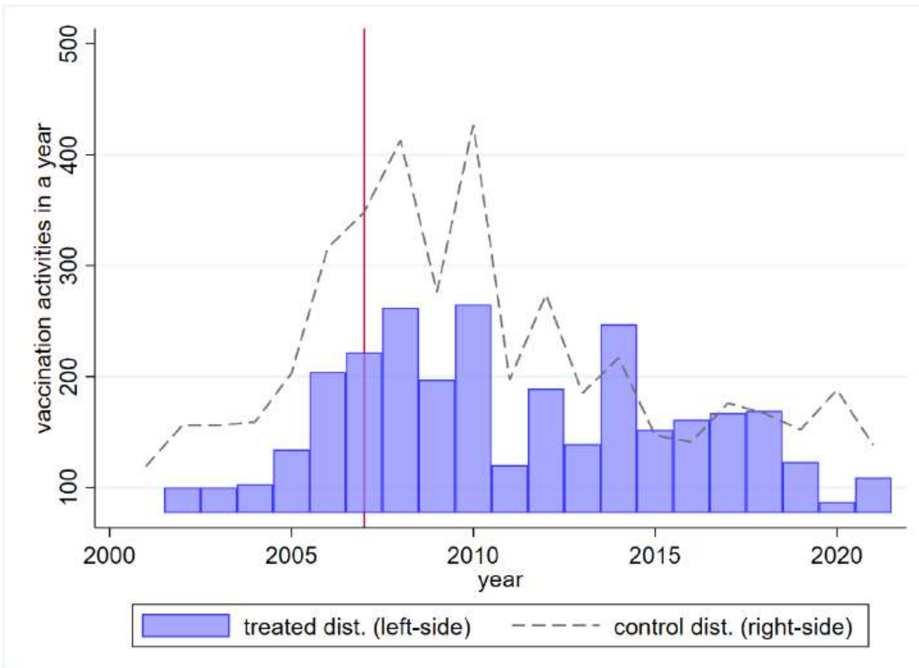


Number of Functional Transit Points bordering NWA

Province	District	Tehsil	Total
FATA	FR Bannu	FR Bannu	6
FATA	FR D I Khan	FR D.I Khan	2
FATA	FR Tank	FR Tank	2
FATA	Kurram Agency	Central Kurram	2
FATA	Kurram Agency	Lower Kurran	4
KP	Bannu	Bannu	14
KP	D I Khan	D I Khan	11
KP	D I Khan	Daraban	1
KP	D I Khan	Kulachi	1
KP	D I Khan	Paharpur	1
KP	D I Khan	Pariva	1
KP	Hangu	Hangu	6
KP	Hangu	Thall	10
KP	Karak	B D Shah	2
KP	Karak	Karak	5
KP	Karak	Takht e Nasrati	1
KP	Lakki Marwat	Lakki Marwat	6
KP	Lakki Marwat	Sarai Naurang	12
KP	Tank	Tank	9
Total			96

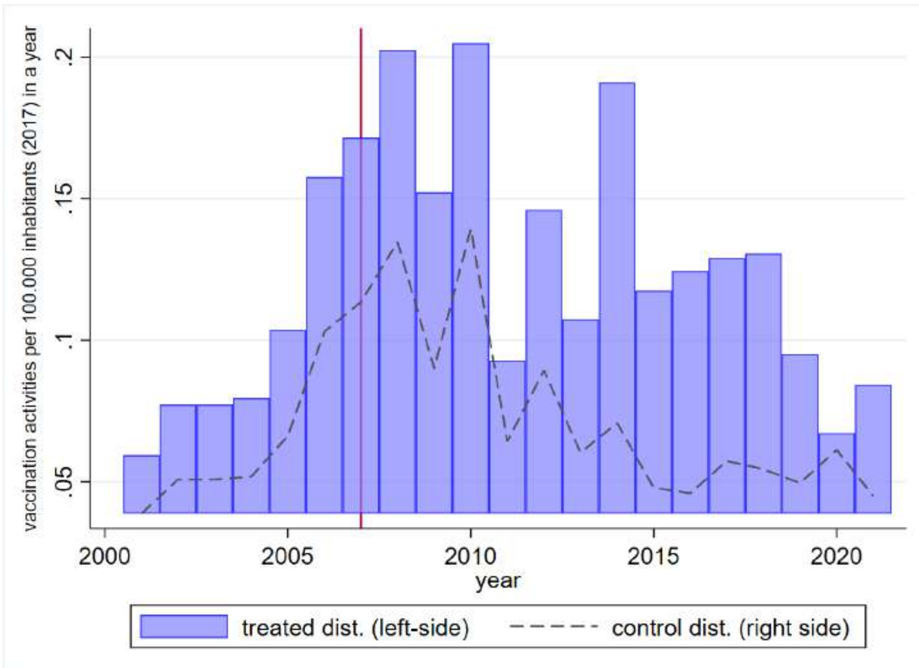
Note: This figure illustrates the functional Permanent Transit Vaccination Points in districts bordering the North Waziristan district. The Pakistan Polio Eradication Programme vaccinates children travelling or on the move through 500 permanent transit points (P.T.P.s) across all major transit points nationwide. These P.T.P.s are set up along country and district borders and other essential transit points such as railway stations, bus stops, and highways. Source. World Health Organization.

Figure A.14: Polio campaigns (2001-2022)



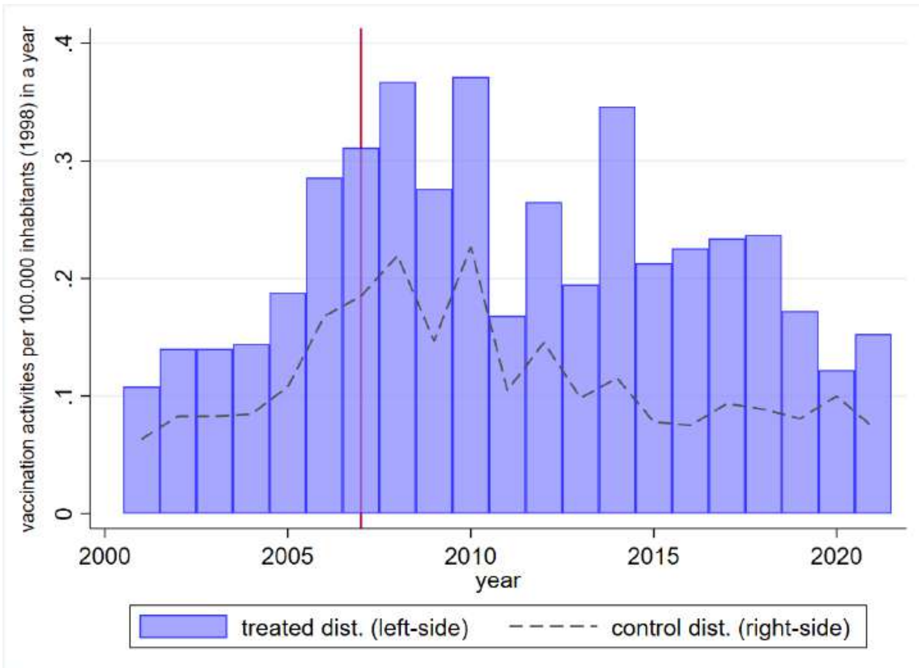
Note: This figure plots the total number of vaccination campaigns against polio in treated and control districts from 2001 to 2022. The blue bars show the campaigns in treated districts, and the grey dashed line in control districts. Treated districts are the host districts and control districts are the non-host districts. Districts whose territory falls within the pre-colonial region of *Pashtunistan* are host districts. Non-host districts are those whose part is outside *Pashtunistan* but adjacent to the historical border. The vertical red line corresponds to December 2007. Source. The Polio Eradication Program from the World Health Organization (WHO).

Figure A.15: Polio campaigns per 100,000 inhabitants (2001-2022)



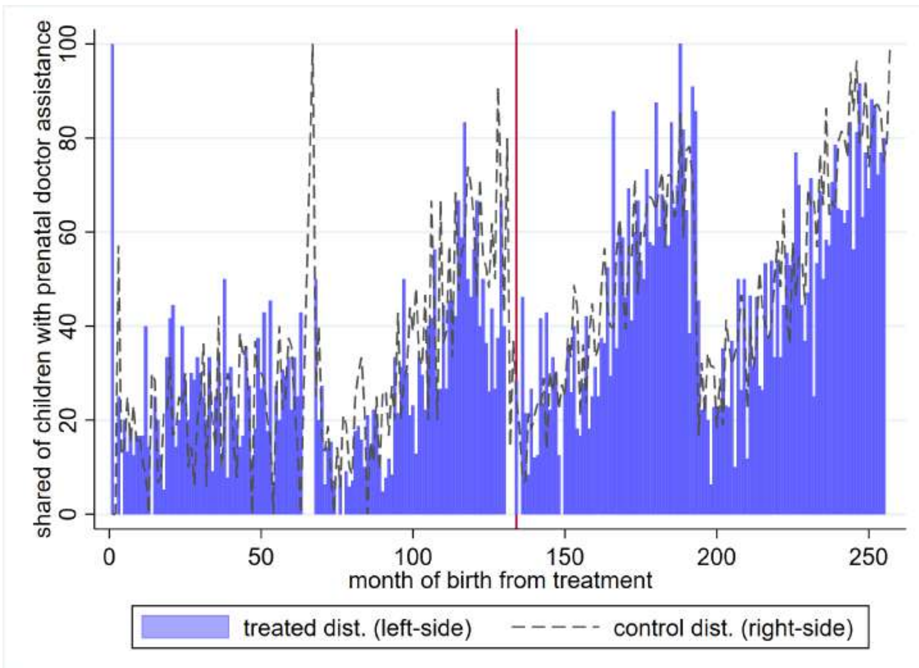
Note: This figure plots the number of vaccination campaigns against polio per 100,000 inhabitants in treated and control districts from 2001 to 2022. I calculate the campaigns per 100,000 inhabitants relative to the population in 2017 from the 2017 population census. The blue bars show the campaigns in treated districts, and the grey dashed line in control districts. Treated districts are the host districts and control districts are the non-host districts. Districts whose territory falls within the pre-colonial region of *Pashtunistan* are host districts. Non-host districts are those whose part is outside *Pashtunistan* but adjacent to the historical border. The vertical red line corresponds to December 2007. Source. The Polio Eradication Program from the World Health Organization (WHO).

Figure A.16: Polio campaigns per 100,000 inhabitants in 1998 (2001-2022)



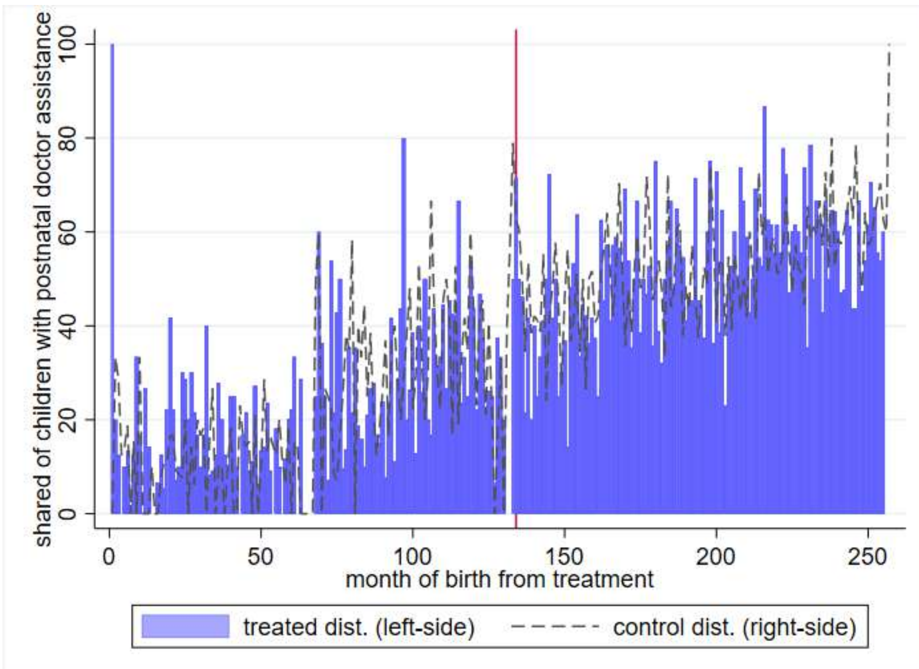
Note: This figure plots the number of vaccination campaigns against polio per 100,000 inhabitants in treated and control districts from 2001 to 2022. I calculate the campaigns per 100,000 inhabitants relative to the population in 1998 from the 1998 population census. The blue bars show the campaigns in treated districts, and the grey dashed line in control districts. Treated districts are the host districts and control districts are the non-host districts. Districts whose territory falls within the pre-colonial region of *Pashtunistan* are host districts. Non-host districts are those whose part is outside *Pashtunistan* but adjacent to the historical border. The vertical red line corresponds to December 2007. Source. The Polio Eradication Program from the World Health Organization (WHO).

Figure A.17: Shared of children with prenatal assistance



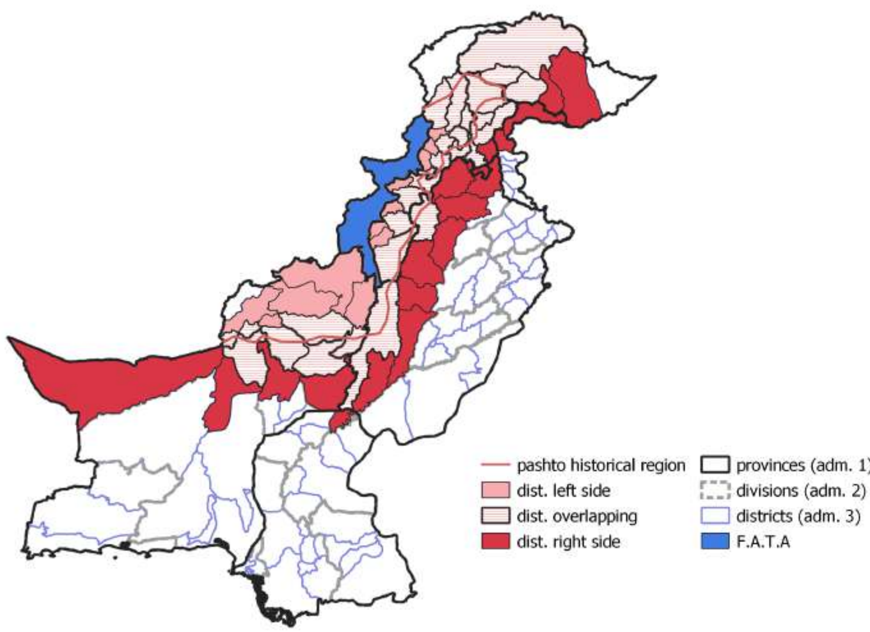
Note: This figure plots the share of children with prenatal assistance in treated and control districts by the birth cohort. Treated districts are the host districts and control districts are the non-host districts. Districts whose territory falls within the pre-colonial region of *Pashtunistan* are host districts. Non-host districts are those whose part is outside *Pashtunistan* but adjacent to the historical border. The vertical red line corresponds to December 2007. Source. Demographic and Health Survey (DHS).

Figure A.18: Shared of children with postnatal assistance



Note: This figure plots the share of children with postnatal assistance in treated and control districts by the birth cohort. Treated districts are the host districts and control districts are the non-host districts. Districts whose territory falls within the pre-colonial region of *Pashtunistan* are host districts. Non-host districts are those whose part is outside *Pashtunistan* but adjacent to the historical border. The vertical red line corresponds to December 2007. Source. Demographic and Health Survey (DHS).

Figure A.19: Alternative treated and control districts



Note: This figure shows the districts partially within *Pashtunistan*. To define them, I use the spatial distribution of districts relative to the pre-colonial region of *Pashtunistan*. The red line corresponds to the *Pashtunistan*'s border. The districts overlapping the border are the red dashed polygons, with only a share of the territory is in *Pashtunistan*. Districts whose territory falls within the pre-colonial region of *Pashtunistan* are host districts. Non-host districts are those whose territory is outside *Pashtunistan* but adjacent to the historical border.

B Additional Tables

Table A.1: Aggregate IDPs by district of origin (2001-2022)

(1)	(2)	(3)	(4)	(5)	(6)	(7)
Province and Division	Position	District	IDP families	IDP individuals	Drones	Total deaths
F.A.T.A	Southern	North Waziristan	108,149	648,894	291	2003
F.A.T.A	Southern	South Waziristan	71,124	426,744	84	678
F.A.T.A	Southern	Largha Shirani	0	0	1	6
F.A.T.A	Northern	Bajaur	72,895	437,370	4	128
F.A.T.A		Khyber	91,689	550,134	6	61
F.A.T.A		Kurram	33,024	198,144	9	83
F.A.T.A		Mohmand	36,759	220,554	0	0
F.A.T.A		Orakzai	35,823	214,938	1	13
N.W.F.P.	Southern	Tank	2,256	13,536	1	5
TOTAL			451,719	2,710,314	396	2,971

Note: This Table shows the aggregate number of internally displaced persons (IDP) from 2001 to 2022 by district of origin. The IDP data source is UNHCR 2022. Columns (6) and (7) present the aggregate number of drones and the number of deaths created from 2001 to 2022 from New-America 2021.

Table A.2: Total IDPs by district of destination (2008-2015)

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Province	District	2008	2010	2011	2012	2013	2014	2015
Khyber Pakhtunkhwa (NWFP)	Adam Khel	1168						
Khyber Pakhtunkhwa (NWFP)	Charsadda	187						
Khyber Pakhtunkhwa (NWFP)	Dir	190						
Khyber Pakhtunkhwa (NWFP)	Hangu	63	5500	5635	5821	5821	6173	2232
Khyber Pakhtunkhwa (NWFP)	Kohat	1237						
Khyber Pakhtunkhwa (NWFP)	Mardan							3504
Khyber Pakhtunkhwa (NWFP)	Nowshera		102127	32499	57771	22076	30352	
Khyber Pakhtunkhwa (NWFP)	Peshawar	21						
Khyber Pakhtunkhwa (NWFP)	Swat	15639						
Khyber Pakhtunkhwa (FATA)	Bajaur	114717						
Khyber Pakhtunkhwa (FATA)	Khyber	110						
Khyber Pakhtunkhwa (FATA)	Kurram	5275						
Khyber Pakhtunkhwa (FATA)	Mohmand	15516						
Khyber Pakhtunkhwa (FATA)	N. Waziristan	11						
Khyber Pakhtunkhwa (FATA)	Orakzai	1632						
Khyber Pakhtunkhwa (FATA)	S. Waziristan	43						
Balochistan	Qilla Abdullah			14397	12438	14978	4166	4632
TOTAL		155809	107627	52531	76030	42875	40691	10368

Note: This Table shows the total number of internally displaced persons (IDP) from 2008 to 2015 by district of destination. The IDP data source is UNHCR 2022. There are no data for 2009 and after 2015.

Table A.3: Number of polio cases per 100,000 inhabitants in 1998

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	polio pop.den.	polio pop.den.	polio pop.den.	polio pop.den.	polio pop.den.	
2007 x Host district	0.011547***	0.011547*	0.012219**	0.012814	0.016859***	
	(0.003284)	(0.004596)	(0.004167)	(0.005705)	(0.002497)	
Observations	7,128	7,128	7,128	6,600	4,740	
Province FE	No	Yes	Yes	Yes	Yes	
Time FE	No	Yes	Yes	Yes	Yes	
Controls	No	No	Yes	Yes	Yes	
Number of provinces		5	5	4	5	

Note: This Table presents the impacts of the IDP inflows on district polio cases per 100,000 inhabitants measured in 1998 in host districts compared to non-host district. I measure the total cases per 100,000 inhabitant, I use the total population at district level from the 1998 Population Census. Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A.4: Pre-treatment characteristics, host vs non-host districts

	(1)	(2)	(3)
	Non-Host	Host	Diff
monthly polio cases	0.026	0.042	0.019
	(0.261)	(0.219)	(0.011)
monthly number of polio campaigns	0.702	0.702	0.000
	(0.457)	(0.457)	(0.000)
nightlight intensity	6.939	7.831	0.495
	(5.437)	(5.841)	(1.060)
monthly number of drones	0.000	0.000	0.000
	(0.457)	(0.457)	(0.000)
monthly number of terrorist attacks	0.070	0.078	0.008
	(0.457)	(0.457)	(0.000)
Household Characteristics in 1998 - Census			
electricity	0.529	0.838	0.278
	(0.018)	(0.126)	(0.000)
piped water	0.152	0.308	-0.003
	(0.091)	(0.078)	(0.000)
own house	0.892	0.807	-0.040
	(0.026)	(0.086)	(0.000)
family size	9.598	11.540	3.317
	(0.795)	(0.871)	(0.000)
number children under 5	0.301	0.301	0.050
	(0.028)	(0.017)	(0.000)
head of hh literate	0.257	0.270	-0.068
	(0.047)	(0.038)	(0.000)
muslim	0.994	0.993	-0.003
	(0.001)	(0.003)	(0.000)
Pashtu mother tongue	0.056	0.816	0.622
	(0.080)	(0.207)	(0.000)
Household Characteristics in 2006 -DHS			
piped water	0.346	0.636	0.104*
	(0.279)	(0.214)	(0.044)
floor	0.379	0.389	0.085
	(0.284)	(0.161)	(0.163)
television	0.310	0.400	0.150***
	(0.254)	(0.152)	(0.018)
watch tv every week	0.206	0.318	0.209***
	(0.208)	(0.095)	(0.031)
radio	0.393	0.555	0.161
	(0.167)	(0.207)	(0.114)
head hh working	0.204	0.127	-0.048**
	(0.136)	(0.115)	(0.015)
number children under 5	2.258	3.032	0.764***
	(0.469)	(0.595)	(0.047)
number members	8.323	12.076	4.068***
	(1.435)	(2.163)	(0.724)
urban	0.288	0.510	0.265*
	(0.215)	(0.306)	(0.105)
Observations	1,596	1,008	2,604

Note: This table reports descriptive statistics for the main variables and sample considered in the baseline analysis. The analysis covers 31 districts from 2001 to 2022 at the monthly level (264 observations per district). For the pre-treatment analysis, I restrict my timeframe from 2001 to 2007. Pre-treatment characteristics are from the 1998 Population Census and the 2006-2007 Demographic and Health Survey (DHS).

Table A.5: Potential conflict confounding effect

	(1)	(2)	(3)	(4)	(5)
VARIABLES	polio	polio	polio	polio	polio
	PANEL A: Northern sample				
2007 x Host district	0.048***	0.048***	0.086***	0.083***	0.129***
	(0.017)	(0.013)	(0.013)	(0.019)	(0.028)
Observations	3,432	3,432	3,432	2,112	2,028
Number of provinces		5	5	2	5
	PANEL B:Southern sample				
2007 x Host district	0.036***	0.036***	0.043***	0.043***	0.037*
	(0.011)	(0.011)	(0.011)	(0.011)	(0.020)
Observations	4,752	4,752	4,752	4,488	3,012
Number of provinces		4	4	4	4
	PANEL C: Terrorist attacks controls				
2007 x Host district	0.025***	0.025***	0.040***	0.044**	0.051***
	(0.009)	(0.006)	(0.007)	(0.010)	(0.013)
Observations	8,184	8,184	8,184	6,600	5,040
Number of provinces		7	7	4	7
Province FE	No	Yes	Yes	Yes	Yes
Time FE	No	Yes	Yes	Yes	Yes
Controls	No	No	Yes	Yes	Yes

Note: This Table presents the results of Table 1 by controlling for conflict. Panel A shows the estimates restricting the sample to Northern district of my main sample. The districts in the North are: Abbottabad, Attok, Chakwal, Charsadda, Hangu, Islamabad, Kargil, Kupwara (Gilgit Wazarat), Malakand P.A., Muzaffarabad, Neelum, Peshawar, and Rawalpindi. Panel B shows the estimates for a sample of Southern districts (Bannu, Bhakkar, Bhittani, Bolan, Chagai, Dera Bugti, Kalat, Kashmore, Khushab, Layyah, Musakhel, Muzaffargarh, Pishin, Qilla Saifullah, Rajan Pur, Tank, Zhob, and Ziarat). Panel C controls for the number of terrorist attacks. Standard errors are clustered at the province level. Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A.6: Potential Afghan refugees effect

	(1)	(2)	(3)	(4)	(5)
VARIABLES	polio	polio	polio	polio	polio
	PANEL A: Control for total refugees in a district				
2007 x Host district	0.025***	0.025**	0.044**	0.048	0.058**
	(0.009)	(0.007)	(0.017)	(0.022)	(0.023)
	PANEL B: Number of refugee camps fixed effects				
2007 x Host district	0.038***	0.038**	0.053**	0.056*	0.066**
	(0.009)	(0.012)	(0.019)	(0.023)	(0.022)
	PANEL C: Number of refugee camps interaction				
2007 x Host district	0.014	0.018**	0.039*	0.045	0.056*
	(0.009)	(0.007)	(0.017)	(0.021)	(0.023)
2007 x Host district X n. camps	0.006***	0.005***	0.003**	0.003	0.002
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Observations	8,184	8,184	8,184	6,600	5,040
Province FE	No	Yes	Yes	Yes	Yes
Time FE	No	Yes	Yes	Yes	Yes
Controls	No	No	Yes	Yes	Yes
Number of provinces		7	7	4	7

Note: This Table presents the results of Table 1 by controlling for the presence of refugees from Afghanistan. Panel A shows the estimates controlling for the number of refugees in a year. Panel B includes number of camps fixed effects. In Panel C, I add an interaction to the number of refugee camps. Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A.7: Potential international migration effects

	(1)	(2)	(3)	(4)	(5)
VARIABLES	polio	polio	polio	polio	polio
2007 x Host district	0.040***	0.040***	0.046**	0.051*	0.049*
	(0.015)	(0.010)	(0.013)	(0.018)	(0.019)
Observations	3,906	3,906	3,906	3,150	1,650
Province FE	No	Yes	Yes	Yes	Yes
Time FE	No	Yes	Yes	Yes	Yes
Controls	No	No	Yes	Yes	Yes
Number of provinces		7	7	4	5

Note: This Table presents the results of Table 1 restricted to the period 2001 to 2011. The number of Pakistani refugees has been relatively constant from 2000 to 2011, with an increase from 2012. Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A.8: Effect of IDP population on vaccination against polio, with province fixed effect

	(1)	(2)	(3)	(4)	(5)
VARIABLES	vaccinated	vaccinated	vaccinated	vaccinated	vaccinated
	PANEL A: Dycotomic treatment				
2007 x Host district	-0.083***	-0.059	-0.031	-0.043	-0.020
	(0.017)	(0.056)	(0.059)	(0.051)	(0.052)
	PANEL B: Continuous treatment - Predicted inflow				
Predicted Inflow	0.021***	-0.004	-0.005	-0.012	-0.011
	(0.005)	(0.015)	(0.016)	(0.017)	(0.017)
Observations	10,608	10,608	10,608	10,563	10,563
Province FE	No	Yes	Yes	Yes	Yes
Time FE	No	Yes	Yes	Yes	Yes
Controls	No	No	Yes	Yes	Yes
Number of provinces		7	7	7	7

Note: This Table presents the results of Table 3 by controlling for province fixed effect. Panel A shows the estimates with the baseline treatment. Panel B shows the results with the predicted inflow. Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A.9: Pre-treatment children characteristics at individual level, host vs non-host districts

	(1) Non-Host	(2) Host	(3) Diff
water piped	0.337	0.615	0.096
	(0.473)	(0.487)	(0.046)
toilet	0.339	0.456	0.248
	(0.473)	(0.498)	(0.158)
floor	0.413	0.390	-0.002
	(0.493)	(0.488)	(0.179)
television	0.359	0.485	0.214***
	(0.480)	(0.500)	(0.041)
watch tv every week	0.249	0.425	0.231***
	(0.432)	(0.495)	(0.011)
radio	0.390	0.488	0.153
	(0.488)	(0.500)	(0.143)
head hh working	0.208	0.072	-0.058*
	(0.406)	(0.258)	(0.025)
number children under 5	2.341	3.073	0.983***
	(1.442)	(2.045)	(0.179)
number members	8.636	11.309	4.316**
	(4.331)	(6.494)	(1.497)
mother education	0.448	0.374	0.002
	(0.846)	(0.785)	(0.089)
diarrhea	0.154	0.146	0.028
	(0.361)	(0.353)	(0.017)
fever	0.260	0.252	0.034
	(0.439)	(0.434)	(0.024)
head hh woman	0.066	0.033	-0.083
	(0.249)	(0.178)	(0.070)
urban	0.326	0.544	0.196***
	(0.469)	(0.498)	(0.031)
girl	0.490	0.471	-0.039**
	(0.500)	(0.499)	(0.014)
Observations	1,596	1,008	2,604

Note: This table reports descriptive statistics for the main variables and sample considered in the baseline analysis. The analysis covers 31 district from 2001 to 2022 at the monthly level (264 observations per district). For the pre-treatment analysis I restrict my timeframe from 2001 to 2007. Pre-treatment characteristics are from the 1991-1992 and 2006-2007 Demographic and Health Survey (DHS).

Table A.10: Effect of IDP population on polio

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	diarrhea	diarrhea	diarrhea	fever	fever	fever
			PANEL A: District fixed-effects			
2007 x Host district	-0.016			-0.148***		
	(0.033)			(0.036)		
Predicted Inflow		-0.020**	-0.020**		-0.024	-0.022
		(0.009)	(0.009)		(0.019)	(0.019)
Predicted Inflow x IDP			-0.027			-0.048***
			(0.020)			(0.016)
Number of districts	31	31	31	31	31	31
			PANEL B: Province fixed-effects			
2007 x Host district	0.016			-0.067**		
	(0.013)			(0.021)		
Predicted Inflow		-0.007	-0.006		-0.009	-0.008
		(0.006)	(0.006)		(0.007)	(0.006)
Predicted Inflow x IDP			-0.028***			-0.048***
			(0.006)			(0.005)
Number of provinces	7	7	7	7	7	7
Observations	10,623	10,623	10,623	10,623	10,623	10,623
Province FE	No	No	No	No	No	No
Time FE	No	No	No	No	No	No
Controls	No	No	No	No	No	No

Note: This Table presents the impacts of the IDP inflows on diarrhea and fever. I use individual level data from the Demographic health surveys. The outcome on diarrhea is equal to one if the child had diarrhea recently, zero otherwise. Fever is 1 if the child had fever the last two weeks. The baseline specification is presented in equation (2). Panel A shows the results with district fixed effects as in equation (2). In panel B, I control for province fixed effects. Standard errors are clustered at the province level. Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A.11: Mistrust on vaccines channel: cases per 100,000 inhabitants

	(1)	(2)	(3)	(4)	(5)
VARIABLES	polio	polio	polio	polio	polio
	PANEL A: Dycotomic treatment, 2001-2011				
2007 x Host district	0.005	0.005*	0.005*	0.005	0.005**
	(0.003)	(0.002)	(0.002)	(0.003)	(0.002)
	PANEL B: Continuos treatment - Predicted Inflow, 2001-2011				
Predicted Inflow	0.000913*	0.000911***	0.000868***	0.000637*	0.000488
	(0.000511)	(0.000167)	(0.000179)	(0.000207)	(0.000258)
Observations	3,240	3,240	3,240	3,000	1,500
Province FE	No	Yes	Yes	Yes	Yes
Time FE	No	Yes	Yes	Yes	Yes
Controls	No	No	Yes	Yes	Yes
Number of province		5	5	4	5
	PANEL C: Girls sample				
VARIABLES	vaccinated	vaccinated	vaccinated	vaccinated	vaccinated
Predicted Inflow	0.006	-0.004	-0.009	-0.004	-0.021
	(0.009)	(0.010)	(0.014)	(0.010)	(0.013)
Observations	3,064	3,064	10,608	3,046	3,046
Province FE	No	Yes	Yes	Yes	Yes
Time FE	No	Yes	Yes	Yes	Yes
Controls	No	No	Yes	Yes	Yes
Number of districts		26	31	26	26

Note: This Table presents the potential effect of vaccine mistrust. In panel A, I repeat the estimates of panel C of Table 1 in a timeframe from 2001 to 2011. In panel B, I repeat the estimates of Panel C of Table 2 restricting my timeframework from 2001 to 2011. Panel C shows the results of of panel A of Table 3 from 2001 to 2011 in a sample of girls. Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A.12: Falsification test: Effects before treatment

	(1)	(2)	(3)	(4)	(5)
	PANEL A: At least one polio case				
VARIABLES	polio	polio	polio	polio	polio
2007 x Host district	-0.017	-0.017	-0.014	-0.012	-0.099
	(0.029)	(0.025)	(0.022)	(0.024)	(0.106)
Observations	2,604	2,604	2,604	2,100	1,620
Number of provinces		7	7	4	5
	PANEL B: Number of polio cases				
VARIABLES	polio cases	polio cases	polio cases	polio cases	polio cases
2007 x Host district	-0.035	-0.035	-0.031	-0.034	-0.137
	(0.034)	(0.031)	(0.027)	(0.032)	(0.124)
Observations	2,604	2,604	2,604	2,100	1,620
Number of provinces		7	7	4	5
	PANEL C: Number of polio cases per 100,000 inhabitants				
VARIABLES	polio pop.den.	polio pop.den.	polio pop.den.	polio pop.den.	polio pop.den.
2007 x Host district	-0.000	-0.000	0.000	0.000	-0.006
	(0.002)	(0.003)	(0.003)	(0.003)	(0.009)
Observations	2,268	2,268	2,268	2,100	1,620
Number of provinces		5	5	4	5
Province FE	No	Yes	Yes	Yes	Yes
Time FE	No	Yes	Yes	Yes	Yes
Controls	No	No	Yes	Yes	Yes

Note: This Table presents the impacts of the IDP inflows on district polio prevalence in host districts compared to non-host district with a placebo treatment timing. The treatment timing starts from September 2001. I use the spatial distribution of districts with respect the pre-colonial region of *Pashtunistan* to define host and non-host districts. Districts whose territory fall within the pre-colonial region of *Pashtunistan* are defined as host districts. Non-host districts are the district whose territory is outside *Pashtunistan*, but are adjacent to the historical border. Observations are at the district and month level from 2001 to 2022. The baseline specification is presented in equation (1). Column (1) presents the results without province, year-month fixed effects and covariates. Column (2) includes province and year-month fixed effects. Column (3) controls for nightlight intensity and total vaccination campaigns. Columns (4) controls for pre-treatment district-covariates (the average number of children under five, the average number of members in a household, and the total share of the literate population from 1973, 1981 and 1998 Population Census). Column (5) control instead for contemporary characteristics (the average number of children under five, the average number of members in a household, shared households with piped water, and shared households with a finished floor). This Table present three different outcomes: at least one case of polio (panel A), total number of polio cases (panel B) and polio cases per 100,000 inhabitants from the 2017 Population Census (panel C). Standard errors are clustered at the province level. Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A.13: Falsification test: non-pashtu districts counterfactual

	(1)	(2)	(3)	(4)	(5)
	PANEL A: At least one polio case				
VARIABLES	polio	polio	polio	polio	polio
2007 x Host district	-0.009*	-0.009	-0.008	-0.004	0.001
	(0.005)	(0.012)	(0.012)	(0.012)	(0.013)
Observations	26,928	26,928	26,928	21,384	19,200
Number of province		8	8	4	8
	PANEL B: Number of polio cases				
VARIABLES	polio cases	polio cases	polio cases	polio cases	polio cases
2007 x Host district	-0.038***	-0.038	-0.036	-0.018	-0.012
	(0.009)	(0.032)	(0.029)	(0.019)	(0.020)
Observations	26,928	26,928	26,928	21,384	19,200
Number of provinces		8	8	4	8
	PANEL C: Number of polio cases per 100,000 inhabitants				
VARIABLES	polio pop.den.	polio pop.den.	polio pop.den.	polio pop.den.	polio pop.den.
2007 x Host district	0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Observations	25,410	25,410	25,410	19,990	18,836
Number of provinces		5	5	4	5
Province FE	No	Yes	Yes	Yes	Yes
Time FE	No	Yes	Yes	Yes	Yes
Controls	No	No	Yes	Yes	Yes

Note: This Table presents the impacts of the IDP inflows on district polio prevalence in host districts compared to non-host district with a placebo counterfactual. The treatment timing starts from 2008. I use the spatial distribution of districts with respect the pre-colonial region of *Pashtunistan* to define host and non-host districts. Treated districts are the district whose territory is outside *Pashtunistan*, but are adjacent to the historical border. They correspond to the non-host districts in equation (1). Control districts are the district whose territory is outside *Pashtunistan*, but are not adjacent to the historical border. These districts are not included in my baseline sample. The baseline specification is presented in equation (1). Column (1) presents the results without province, year-month fixed effects and covariates. Column (2) includes province and year-month fixed effects. Column (3) controls for nightlight intensity and total vaccination campaigns. Columns (4) controls for pre-treatment district-covariates (the average number of children under five, the average number of members in a household, and the total share of the literate population from 1973, 1981 and 1998 Population Census). Column (5) control instead for contemporary characteristics (the average number of children under five, the average number of members in a household, shared households with piped water, and shared households with a finished floor). This Table present three different outcomes: at least one case of polio (panel A), total number of polio cases (panel B) and polio cases per 100,000 inhabitants from the 2017 Population Census (panel C). Standard errors are clustered at the province level. Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A.14: Treatment definition: districts entirely or partially in *Pashtunistan*

	(1)	(2)	(3)	(4)	(5)
	PANEL A: At least one polio case				
VARIABLES	polio	polio	polio	polio	polio
2007 x Host district	0.021***	0.021**	0.031**	0.030*	0.033**
	(0.005)	(0.006)	(0.009)	(0.010)	(0.012)
Observations	15,312	15,312	15,312	12,672	9,504
Number of provinces		7	7	4	7
	PANEL B: Number of polio cases				
VARIABLES	polio cases	polio cases	polio cases	polio cases	polio cases
2007 x Host district	0.044***	0.044**	0.057**	0.060*	0.063**
	(0.010)	(0.013)	(0.017)	(0.020)	(0.021)
Observations	15,312	15,312	15,312	12,672	9,504
Number of provinces		7	7	4	7
	PANEL C: Number of polio cases per 100,000 inhabitants				
VARIABLES	polio pop.den.	polio pop.den.	polio pop.den.	polio pop.den.	polio pop.den.
2007 x Host district	0.007***	0.007*	0.007**	0.007	0.010***
	(0.002)	(0.003)	(0.002)	(0.003)	(0.001)
Observations	7,128	7,128	7,128	6,600	4,740
Number of provinces		5	5	4	5
Province FE	No	Yes	Yes	Yes	Yes
Time FE	No	Yes	Yes	Yes	Yes
Controls	No	No	Yes	Yes	Yes

Note: This Table presents the impacts of the IDP inflows on district polio prevalence in an alternative definition of host districts compared to baseline definition of non-host district. Host districts are the district whose territory is entirely or partially inside *Pashtunistan*. The treatment timing starts from 2008. Observations are at the district and month level from 2001 to 2022. The baseline specification is presented in equation (1). Column (1) presents the results without province, year-month fixed effects and covariates. Column (2) includes province and year-month fixed effects. Column (3) controls for nightlight intensity and total vaccination campaigns. Columns (4) controls for pre-treatment district-covariates (the average number of children under five, the average number of members in a household, and the total share of the literate population from 1973, 1981 and 1998 Population Census). Column (5) control instead for contemporary characteristics (the average number of children under five, the average number of members in a household, shared households with piped water, and shared households with a finished floor). This Table present three different outcomes: at least one case of polio (panel A), total number of polio cases (panel B) and polio cases per 100,000 inhabitants from the 2017 Population Census (panel C). Standard errors are clustered at the province level. Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A.15: Treatment definition: districts partially in *Pashtunistan*

	(1)	(2)	(3)	(4)	(5)
			PANEL A: At least one polio case		
VARIABLES	polio	polio	polio	polio	polio
2007 x Host district	0.013**	0.013**	0.001	0.017	0.019*
	(0.006)	(0.004)	(0.008)	(0.008)	(0.008)
Observations	12,144	12,144	15,312	9,768	7,452
Number of provinces		7	7	4	7
			PANEL B: Number of polio cases		
VARIABLES	polio cases	polio cases	polio cases	polio cases	polio cases
2007 x Host district	0.027**	0.027**	-0.001	0.036**	0.041**
	(0.012)	(0.008)	(0.016)	(0.011)	(0.013)
Observations	12,144	12,144	15,312	9,768	7,452
Number of provinces		7	7	4	7
			PANEL C: Number of polio cases per 100,000 inhabitants		
VARIABLES	polio pop.den.	polio pop.den.	polio pop.den.	polio pop.den.	polio pop.den.
2007 x Host district	0.001575*	0.001100*	0.000512	0.001175**	0.002044**
	(0.001872)	(0.002769)	(0.002445)	(0.003326)	(0.001200)
Observations	7,128	7,128	7,128	6,600	4,740
Number of provinces		5	5	4	5
Province FE	No	Yes	Yes	Yes	Yes
Time FE	No	Yes	Yes	Yes	Yes
Controls	No	No	Yes	Yes	Yes

Note: This Table presents the impacts of the IDP inflows on district polio prevalence in an alternative defnition of host districts compared to baseline definition of non-host district. Host districts are the district whose territory is partially inside *Pashtunistan*. The treatment timing starts from 2008. Observations are at the district and month level from 2001 to 2022. The baseline specification is presented in equation (1). Column (1) presents the results without province, year-month fixed effects and covariates. Column (2) includes province and year-month fixed effects. Column (3) controls for nightlight intensity and total vaccination campaigns. Columns (4) controls for pre-treatment district-covariates (the average number of children under five, the average number of members in a household, and the total share of the literate population from 1973, 1981 and 1998 Population Census). Column (5) control instead for contemporary characteristics (the average number of children under five, the average number of members in a household, shared households with piped water, and shared households with a finished floor). This Table present three different outcomes: at least one case of polio (panel A), total number of polio cases (panel B) and polio cases per 100,000 inhabitants from the 2017 Population Census (panel C). Standard errors are clustered at the province level. Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A.16: Counterfactual definition: districts partially in *Pashtunistan*

	(1)	(2)	(3)	(4)	(5)
	PANEL A: At least one polio case				
VARIABLES	polio	polio	polio	polio	polio
2007 x Host district	0.025***	0.025*	0.031*	0.036	0.050**
	(0.009)	(0.010)	(0.012)	(0.016)	(0.011)
Observations	10,296	10,296	10,296	8,976	6,516
Number of provinces		4	4	3	4
	PANEL B: Number of polio cases				
VARIABLES	polio cases	polio cases	polio cases	polio cases	polio cases
2007 x Host district	0.054***	0.054	0.063	0.071	0.077*
	(0.016)	(0.029)	(0.032)	(0.042)	(0.032)
Observations	10,296	10,296	10,296	8,976	6,516
Number of provinces		4	4	3	4
	PANEL C: Number of polio cases per 100,000 inhabitants				
VARIABLES	polio pop.den.	polio pop.den.	polio pop.den.	polio pop.den.	polio pop.den.
2007 x Host district	0.005878***	0.005874**	0.006291*	0.007001	0.009333**
	(0.002)	(0.003)	(0.002)	(0.003)	(0.001)
Observations	9,813	9,813	9,813	8,976	6,202
Number of provinces		5	5	4	5
Province FE	No	Yes	Yes	Yes	Yes
Time FE	No	Yes	Yes	Yes	Yes
Controls	No	No	Yes	Yes	Yes

Note: This Table presents the impacts of the IDP inflows on district polio prevalence in the baseline defnition of host districts compared to an alternative definition of non-host district. Non-host districts are the district whose territory is partially inside *Pashtunistan*. The treatment timing starts from 2008. Observations are at the district and month level from 2001 to 2022. The baseline specification is presented in equation (1). Column (1) presents the results without province, year-month fixed effects and covariates. Column (2) includes province and year-month fixed effects. Column (3) controls for nightlight intensity and total vaccination campaigns. Columns (4) controls for pre-treatment district-covariates (the average number of children under five, the average number of members in a household, and the total share of the literate population from 1973, 1981 and 1998 Population Census). Column (5) control instead for contemporary characteristics (the average number of children under five, the average number of members in a household, shared households with piped water, and shared households with a finished floor). This Table present three different outcomes: at least one case of polio (panel A), total number of polio cases (panel B) and polio cases per 100,000 inhabitants from the 2017 Population Census (panel C). Standard errors are clustered at the province level. Robust standard errors are in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table A.17: Counterfactual definition: non-Pashtu districts not adjacent to *Pashtunistan* border

	(1)	(2)	(3)	(4)	(5)
	PANEL A: At least one polio case				
VARIABLES	polio	polio	polio	polio	polio
2007 x Host district	0.029***	0.029*	0.030*	0.038*	0.060***
	(0.009)	(0.014)	(0.013)	(0.015)	(0.008)
Observations	25,080	25,080	25,080	20,592	15,024
Number of provinces		7	7	4	7
	PANEL B: Number of polio cases				
VARIABLES	polio cases	polio cases	polio cases	polio cases	polio cases
2007 x Host district	0.043***	0.043	0.044	0.072	0.102***
	(0.014)	(0.041)	(0.040)	(0.039)	(0.020)
Observations	25,080	25,080	25,080	20,592	15,024
Number of provinces		7	7	4	7
	PANEL C: Number of polio cases per 100,000 inhabitants				
VARIABLES	polio pop.den.	polio pop.den.	polio pop.den.	polio pop.den.	polio pop.den.
2007 x Host district	0.004565**	0.004002*	0.005269*	0.007175*	0.009276*
	(0.001872)	(0.002769)	(0.002445)	(0.003326)	(0.001200)
Observations	24,802	24,802	24,802	20,592	14,871
Number of provinces		5	5	4	5
Province FE	No	Yes	Yes	Yes	Yes
Time FE	No	Yes	Yes	Yes	Yes
Controls	No	No	Yes	Yes	Yes

Note: This Table presents the impacts of the IDP inflows on district polio prevalence in the baseline defnition of host districts compared to an alternative definition of non-host district. Non-host districts are the district outside *Pashtunistan*, and non-adjacent to *Pashtunistan* border. The treatment timing starts from 2008. Observations are at the district and month level from 2001 to 2022. The baseline specification is presented in equation (1). Column (1) presents the results without province, year-month fixed effects and covariates. Column (2) includes province and year-month fixed effects. Column (3) controls for nightlight intensity and total vaccination campaigns. Columns (4) controls for pre-treatment district-covariates (the average number of children under five, the average number of members in a household, and the total share of the literate population from 1973, 1981 and 1998 Population Census). Column (5) control instead for contemporary characteristics (the average number of children under five, the average number of members in a household, shared households with piped water, and shared households with a finished floor). This Table present three different outcomes: at least one case of polio (panel A), total number of polio cases (panel B) and polio cases per 100,000 inhabitants from the 2017 Population Census (panel C). Standard errors are clustered at the province level. Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A.18: Alternative specifications

	(1)	(2)	(3)	(4)	(5)
VARIABLES	polio	polio	polio	polio	polio
	PANEL A: Division fixed effect				
2007 x Host district	0.038***	0.038**	0.053*	0.060*	0.062
	(0.009)	(0.017)	(0.028)	(0.031)	(0.039)
Observations	8,184	8,184	8,184	6,600	5,040
Number of divisions		19	19	15	19
Division FE	No	Yes	Yes	Yes	Yes
Time FE	No	Yes	Yes	Yes	Yes
	PANEL B: Province and Year fixed effects				
2007 x Host district	0.038***	0.038**	0.053**	0.055*	0.067**
	(0.009)	(0.012)	(0.018)	(0.022)	(0.022)
Observations	8,184	8,184	8,184	6,600	5,040
Number of provinces		7	7	4	7
Province FE	No	Yes	Yes	Yes	Yes
Time FE	No	Yes	Yes	Yes	Yes
	PANEL C: Province fixed effects				
2007 x Host district	0.038***	0.038**	0.052**	0.055*	0.068**
	(0.009)	(0.012)	(0.019)	(0.023)	(0.022)
Observations	8,184	8,184	8,184	6,600	5,040
Number of provinces		7	7	4	7
Province FE	No	Yes	Yes	Yes	Yes
	PANEL D: No clusters				
2007 x Host district	0.038***	0.038***	0.053***	0.056***	0.067***
	(0.009)	(0.008)	(0.008)	(0.010)	(0.016)
Observations	8,184	8,184	8,184	6,600	5,040
Number of provinces		7	7	4	7
Province FE	No	Yes	Yes	Yes	Yes
Time FE	No	Yes	Yes	Yes	Yes
Controls	No	No	Yes	Yes	Yes

Note: This Table presents the results of Table 1 for different specifications. Panel A shows the estimates controlling division fixed effect instead of province fixed effect. Panel B control only for province and year fixed effects. In Panel C, I only control for province fixed effect. Panel D shows the estimates without clustering the error terms. Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table A.19: Potential reverse causality

	(1)	(2)	(3)	(4)	(5)
VARIABLES	predicted inflow	predicted inflow	predicted inflow	predicted inflow	predicted inflow
yearly polio cases	0.031668***	0.014447	0.011966	-0.007485	0.010872
	(0.003580)	(0.017856)	(0.016618)	(0.003425)	(0.016732)
Observations	8,184	8,184	8,184	6,600	5,040
Province FE	No	Yes	Yes	Yes	Yes
Time FE	No	Yes	Yes	Yes	Yes
Controls	No	No	Yes	Yes	Yes
Number of provinces		7	7	4	7

Note: This Table presents the estimates of the yearly polio cases from 2001 to 2007 on the predicted inflow. This estimation allows me to evaluate the potential reverse causality of the number of cases on the decision on where to migrate. Robust standard errors are in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$