

Online Appendix to "Stimulant or depressant? Resource-related income shocks and conflict"

Review of Economics and Statistics

Kai Gehring*

Sarah Langlotz[†]

Stefan Kienberger[‡]

31st August 2023

Contents

A	Definition and sources of the variables	3
B	Descriptive statistics	14
C	Geographical overview	17
D	Identification using complement prices	19
D.1	Theoretical predictions	19
D.2	Simulation	22
D.3	Derivations and other results	28
E	Further results	31
E.1	Changes in international prices and local revenues	31
E.2	Instrumental variable specifications	31
E.3	Aggregate results	35
E.4	Counter-factual when freezing prices	36
E.5	Opportunity costs and shifts in conflict tactics	37

*Corresponding author: University of Zurich, e-mail: mail@kai-gehring.net

[†]University of Goettingen, e-mail: sarah.langlotz@uni-goettingen.de

[‡]University of Salzburg, e-mail: stefan.kienberger@sbg.ac.at

CONTENTS	2
F Sensitivity analysis	41
G Maps	68
H Data coding and map generation	73

A. Definition and sources of the variables

Any Processing Lab: We count all types of heroin laboratories. This variable takes on the value 1 if there is at least one lab in a district i , and 0 otherwise. As described in Appendix H, we geo-reference maps from UNODC reports regarding drug markets, labs, and trafficking routes, assign coordinates to the labs, and later compute district averages. Source: UNODC (2006/07, 2014, 2016).

Any Military Base: This variable takes on the value 1 if there is at least one open military base in a district i in year t , and 0 otherwise. The approach is described in detail in Appendix H. Note that we are most likely not capturing all existing locations, as we did not receive the exact information about opening and closing for all military bases. Opening and closing dates were coded with the available information; if there was no information about shutting down a base we assume it is still active. Source: For the more well-known bases, we use Wikipedia's GeoHack program; for the less well-documented bases, we use Wikimapia and Google Maps satellite data.

Battle-Related Deaths (BRD): This variable measures the best (most likely) estimate of total fatalities resulting from an event, with an event being defined as [an] incident where armed force was [used] by an organized actor against another organized actor, or against civilians, resulting in at least 1 direct death at a specific location and a specific date." A direct death is defined as a death relating to either combat between warring parties or violence against civilians"(Sundberg & Melander, 2013; Croicu & Sundberg, 2015). These battle-related deaths include dead civilians and deaths of persons of unknown status. Note that the Uppsala Conflict Data Program Georeferenced Event Dataset (UCDP GED) only includes BRD of events that belong to a dyad ("two conflicting primary parties or party killing unarmed civilians") that reached in total at least 25 BRD within one year. If the dyad generated events with less than 25 BRD in the previous or subsequent years, they are still counted if the dyad had reached the 25 BRD threshold in another year. At the country level, the thresholds 25 and 1000 are common. Berman & Couttenier (2015), in contrast, use a one-BRD threshold at a small grid cell level. While districts are in between these extremes with regard to size, a one-BRD threshold might suffer from misreporting and falsely coding conflict. We construct a continuous measure (log

of BRD) and binary outcomes from all BRD of any party or any type of violence (state-based, non-state or one-sided violence). To capture the lowest level of conflict in a binary measure, we classify a district-year observation with at least five BRD *small conflict*. We then increase the threshold to 10 for the next level of conflict intensity (*low conflict*). In analogy to the threshold used in macro level analyses, we call a district-year observation *conflict* if there are more than 25 BRD. At the top, we take a threshold of 100 BRD for the most severe level of violence what we call *war*. Since UCDP GED provides information on the parties and the type of violence we also construct specific outcome measures according to those categories. Besides different measures of incidence, we also construct measures on onset and ending. We define conflict onset as the incidence of a conflict in a district, where there was no conflict in the previous year ($Conflict_{i,t} = 1 | Conflict_{i,t-1} = 0$). Years of ongoing conflict are set to missing. In analogy, a conflict ending is defined when conflict persisted in the previous year but not in the current year ($Conflict_{i,t} = 0 | Conflict_{i,t-1} = 1$). We also set the ending variable missing for observations which have been at peace in the previous year and remained in peace in the current year, following the standards in the literature. Source: UCDP GED (Sundberg & Melander, 2013; Croicu & Sundberg, 2015).

Calorie Intake: We use a questionnaire in which women self report amounts, frequencies, and sources of a large set of food items, to construct measures on calorie intake and food insecurity. We multiply amounts consumed with kcal values for that food item to get total household calorie intake. Total household daily calorie intake is divided by the number of members that were resident and ate at least dinner regularly in the household during the last seven days to get per capita measures. Source: For kcal values, we use the CSO & The World Bank (2011). The questionnaire responses are from the NRVA women's questionnaire (CSO, 2005, 2007/08, 2011/12).

Consumer Price Index (CPI): Source: For the Euro area (19 countries), we draw data from the OECD (2016); for the remaining countries (2010 = 100), we use the World Bank (2016).

Dietary Diversity: This variable varies between 0 and 8, with eight indicating a high food diversity. According to Wiesmann *et al.* (2009, p. 5) "Dietary diversity is defined as the

number of different foods or food groups eaten over a reference time period, which in my case is one week, not regarding the frequency of consumption.” We classify the different food items from the survey into eight food groups as explained in Wiesmann *et al.* (2009). These groups are staples, pulses, vegetables, fruit, meat/fish, milk/dairy, sugar, and oil/fat. Source: NRVA (CSO, 2005,2007/08,2011/12).

Distance/Proximity/Travel Time to Kabul (capital) and Kandahar, Kunduz, Jalalabad, Hirat, and Mazari Sharif (next five largest cities): For the proximity to Kabul and other main cities we define binary indicators for the distance being smaller than 75 km (1 if < 75) or smaller than 100 km (1 if < 100). In analogy to these categories, we construct indicators for the travel time to Kabul or one of the other main cities falling below 2 or 3 h. We use the shapefiles provided by the Afghan statistical authority on the 398 Afghan districts. Note that the shapefiles available at www.gadm.org do not reflect the current status of administrative division in Afghanistan, and instead we use the one from Empirical Studies of Conflict (ESOC Princeton, <https://esoc.princeton.edu/data/administrative-boundaries-398-districts>). To compute the distances, we first create the centroid of each district polygon. To compute road distances we combined road shapefiles from the official Afghan authorities with street maps from open street map, which were improved by voluntary contributors to close gaps in the official maps. 3D-distances were computed using elevation data from the US Geological Survey (https://www.usgs.gov/centers/eros/science/usgs-eros-archive-digital-elevation-global-30-arc-second-elevation-gtopo30?qt-science_center_objects=0#qt-science_center_objects, accessed July 9, 2018). We add the elevation information to the shapefile containing the roads, and then compute and save three-dimensional distances. We then use the network analyst in ArcGIS to set up a network between all district centroids, clipping centroids that do not overlap with a street in that district that is closest with regard to the as-the-bird-flies distance. Then, we compute the most efficient routes using road distances in two- and three-dimensions. The distances are saved in a matrix and exported in a table that is further processed in Stata. For the variable “distance to other main cities” we use the minimum distance to any of the five cities. For travel time we use the distinction of roads in three classes (motorways, rural, urban), and assign commonly used values for average traveling speed for that road type based on three

sources. Sources: The first source is UNESCAP

(<http://www.unescap.org/sites/default/files/2.4.Afghanistan.pdf>, p. 14, last accessed August 28, 2019) which assumes that the speed on motorways is 90 km/h and on urban roads 50 km/h.

The second source is IRU (<https://www.iru.org/apps/infocentre-item-action?id=560&lang=en>) which states no limits except for urban areas with 50 km/h. The 3rd source is WHO

(<http://apps.who.int/gho/data/view.main.51421>, last accessed August 28, 2019) reporting 90 km/h for rural. We choose the following average traveling speeds, assuming that no strictly enforced limits and little traffic on motorways (120 km/h), and accounting for some (90km/h-10km/h) and moderate traffic in cities (50-20 km/h). Thus our main choice is the following. Motorways: 120 km/h, rural: 80 km/h, urban: 30 km/h. These choices are not perfect, but we verify that our results hold with other variations as well.

Drug Prices (International): Variables are normalized so that prices vary between 0 and 1. We use data on average prices per gram across all available countries in Europe for the following drugs: amphetamines, cocaine, ecstasy, heroin (brown). To construct the average price of alternative drugs we use a mean of the three stimulant drugs amphetamines, cocaine, and ecstasy. For the analysis we convert all drug prices into constant 2010 euros per gram. We then normalize the prices by using a linear min-max function such that all prices vary between 0 and 1. Some regressions use the logarithm of prices using the transformation of $\log(\text{price}+0.01)$. Source: European Monitoring Center for Drugs and Drug Addiction (EMCDDA).

Economically Improved: This variable refers to the question “How do you compare the overall economic situation of the household with 1 year ago?” A value of 1 indicates much worse, 2 slightly worse, 3 same, 4 slightly better, and 5 much better. This is a self-reported measure. Source: NRVA (CSO, 2005, 2007/08, 2011/12).

Ethnic Groups: We record the presence of ethnic groups on a district level based on the Ethnologue (Gordon, 2005) and the the geo-referencing of ethnic groups’ (GREG, Weidmann *et al.*, 2010) datasets. We use the GIS-coordinates of all ethnic groups. Both datasets are based on the historical distribution of ethnic groups or language diffusion. This is an advantage as it ensures that the boundaries are not endogenous to changes during our period of observation. It is partly a disadvantage if groups and countries changed over time. In Afghanistan, the country

boundary did not change. Ethnic group populations certainly change to some degree over time, so that all variables more precisely capture the historic homelands of ethnic groups rather than the current settlement areas. An alternative definition relies on native languages (including Pashto) present in a district using the NRVA 2003 household survey, which is available for a subset of districts. An NRVA-based measure of district with a Pashtun majority correlates more strongly with the Ethnologue measures.

Ethnically Mixed: We construct two measures of whether a district is ethnically mixed. The first variable is the number of ethnic groups; the second is a binary indicator, which takes a value of 1 if the number of ethnic groups is larger than 1, and 0 otherwise. For more extensive methodology, see Ethnic Groups. Source: GREG dataset (Weidmann *et al.* , 2010).

Ethnic Trafficking Route: The variable takes on the value of 1 if there is a potential trafficking route leading from a district to at least one unofficial border crossing point without crossing the ethnic homeland of another group. The underlying intuition is that trafficking is cheaper and significantly easier to conduct, and the accruing additional profits are higher, if there is no need to cross the area of other ethnic groups to transport over the border. Source: For data on unofficial border crossings, we used the UNODC; for information about the homelands of ethnic groups, we used the (GREG) dataset (Weidmann *et al.* , 2010).

Food Expenditures (Paasche/Laspeyres): Precise food amounts were merged with local prices to estimate household food expenditure. We show three food expenditure measures, which are all measured in constant 2011 prices, i.e., prices of the 2011/12 survey wave. Only food items that appear in all three waves are included to build the measure. The first measure “Food Exp. 2011 Prices” does not account for spatial price differences. “Food Exp. 2011 Prices, Paasche” and “Food Exp. 2011 Prices, Laspeyres” adjust for spatial price differences, since households in different districts face different prices. Missing values of district prices are replaced by the province median, which in case of missing values has been replaced by the national median price. For close to all reported food items, prices have been given in the district questionnaire. Prices vary at the district level. Following the literature, we include food items from all possible sources, i.e., purchased food or food in form of gifts etc. Information on food and drinks consumed outside the house (from the male survey section) are also included

in the total food expenditure measures (adjusted for inflation and regional price differences depending on the measure). Expenditures are measured in per capita terms by dividing the total household food expenditure with the number of households (resident and ate at least dinner regularly in the household during the last seven days). We use the section on food consumption from the NRVA women's questionnaire as this section offers precise amounts per food item. Source: NRVA women's and male's questionnaire and district questionnaire (CSO, 2005, 2007/08, 2011/12).

Inflation, GDP Deflator: We use a GDP deflator for the United States with 2010 as the base year. Source: World Bank (2016).

Insecurity/Violence Shock: The share of sampled households per district that have experienced a shock due to insecurity/violence. At the household level, the variable takes on the value of 1 if the household has experienced an insecurity/violence shock. Source: NRVA survey (CSO, 2005, 2007/08, 2011/12).

Legal Opioids: Since most single publications do not cover our whole sample period, we want to cross-verify the numbers using a variety of sources. Source: A main source is the US CDC Public Health surveillance report 2017 (<https://stacks.cdc.gov/view/cdc/47832>, last accessed August 28, 2019). Other important sources were Manchikanti *et al.* (2012); Kenan & Mack (2012); Dart *et al.* (2015).

Local Opium Price: We utilize reports of (monthly) province-level dry opium prices by farmers and by traders as well as country-wide yearly data on fresh opium farm-gate prices weighted by regional production. The province-level opium prices of farmers and traders are highly correlated, with a correlation coefficient close to 1 (0.998). The correlation between the country level farm-gate price and the province-level farm-gate price is 0.66, significant at the 1%-level. While the province-level prices are only available from 2006 to 2013 and for a subset of provinces, they are still very helpful in identifying whether international prices are correlated with local prices. We use the country-wide yearly data on fresh opium farm-gate prices in Afghanistan interacted with the suitability as one proxy for opium profitability in our regressions in Table 1, panel A. Source: Annual Afghanistan Opium Price Monitoring reports (UNODC).

Luminosity: We use this variable as a proxy for GDP and development (Henderson *et al.* , 2012). The yearly satellite data are cloud-free composites made using all the available smooth resolution data for calendar years. The products are 30 arc second grids, spanning -180 to 180 degrees longitude and -65 to 75 degrees latitude. A number of constraints are used to select the highest quality data for entry into the composites: Data are from the center half of the 3000 km wide OLS swaths. Lights in the center half have better geolocation, are smaller, and have more consistent radiometry. Sunlit data and glare are excluded based on the solar elevation angle, Moonlit data based on a calculation of lunar illuminance. Observations with clouds are excluded based on clouds identified with the OLS thermal band data and NCEP surface temperature grids. Lighting features from the aurora have been excluded in the northern hemisphere on an orbit-by-orbit manner using visual inspection. Source: Version 4 DMSP-OLS nighttime lights time series, National Oceanic and Atmospheric Administration-National Geophysical Data Center (NOAA/NGDC, <https://www.ngdc.noaa.gov>, last accessed August 28, 2019). We take the logarithm.

Markets (Major/Sub) and Sum of all Markets: The first variable takes on the value 1 if there is at least one major or sub-market in district i , and 0 otherwise. The second variables counts the sum of all opium markets in a district (both sub and major). Source: UNODC reports on drug markets, labs, and trafficking routes (e.g., UNODC 2006/07, 2014, 2016).

Market Access: Market access for a district i is computed as $MA_i = \sum_{j=1}^N dist_{i,j}^{-\theta} W_j$. W_j is the importance of district j proxied using either the number of opium markets or mean luminosity (or population). $dist_{i,j}$ are the distances between the district and the other districts and θ is the factor discounting other districts that are further way. We use a factor of 1, as in Donaldson & Hornbeck (2016). To take account of the topography and mountainous terrain in Afghanistan, we compute distances using the two-dimensional road network (Market Access 2D) as well as a three-dimensional road network when adjusting for elevation (Market Access 3D).

Mixed/Taliban Territory 1996: The binary indicator on Taliban Territory that we create takes on the value 1 if a district belongs to the territory that was occupied or under the control of the Taliban in 1996, and 0 otherwise. A second indicator (Taliban Territory 1996 - No North)

takes on a value of 1 if the district is exclusively occupied by the Taliban and is characterized by no presence of the Northern Alliance. We use an existing map which indicates the territory of the Taliban in 1996 as well as the territory of other major groups of the Northern Alliance (Dschunbisch-o Islami, Dschamiat-i Islami, Hizb-i Wahdat). We geo-referenced the map and aligned it with the district boundaries; in many cases, the division was quite clearly aligned or overlapping with a district boundary, in the other cases we chose the closest district boundary. We classify a district as a Mixed Territory if it is part of the Taliban 1996 territory and part of the territory of any of the three groups belonging to the Northern Alliance. Source: The map is from Dorronsoro (2005), and more details can be found in Giustozzi (2009).

Opium Cultivation and Revenues: This variable measures opium cultivation in hectares. Data at the district level is an estimate from the data at the province level. We use logged values for opium cultivation and for revenues. From opium cultivation and the respective yields we were able to calculate actual opium production at the district-year level. We also constructed opium revenues by multiplying opium production in kg with the fresh opium farm-gate prices at harvest time in constant 2010 EU/kg. Source: Annual Opium Poppy Survey (UNDCP, 2000) and Afghanistan Opium Survey (UNODC, 2001-2014).

Opium Suitability: This is an index with possible values ranging between 0 and 1 which acts as a proxy for potential of opium production based on exogenous underlying information about land cover, water availability, climatic suitability, and soil suitability. The environmental as well as climatic suitability to cultivate opium poppy (*Papaver somniferum*) is characterized by different factors such as the prevailing physio-geographical and climatic characteristics using climatic suitability based on the EcoCrop model from Hijmans *et al.* (2001). The factor determined to be most important by experts is land cover (S1, 0.41 – the sum of the weights equals 1.0), followed by water availability (S2, 0.28) and climatic conditions (S3, 0.21) respectively. This is in line with additional studies previously carried out by UNODC and described in the World Drug Report (2011) for Myanmar. The data and the index itself was modeled on a 1km^2 resolution and then aggregated to the district units by an area weighted mean approach. The original indicator values were normalized using a linear min–max function between a possible value range of 0 and 100 to allow for comparison and aggregation. Only the land cover indicator

was normalized integrating expert judgments through an Analytical Hierarchy Process (AHP) approach. The four indicators were then subsequently aggregated applying weighted means (weights were verified through expert consultations building on the AHP method). None of the input factors constituting the index is itself to a major degree affected by conflict, which is the outcome variable. Consequently, the index values by district can be considered as exogenously given.

We weight the opium and wheat suitabilities with the (lagged) population distribution within the districts. This is helpful as, for instance, the south features large desert areas and at the same time concentrated areas with dense population, and accounting for the suitability in uninhabited desert areas might be misleading (although our results are not significantly affected by this choice).

Source: The index was developed in the context of a study in collaboration with UNODC; and is described in detail in a publication in a geographical science journal (Kienberger *et al.* , 2017).

Pashtun: Our binary indicator takes on the value 1 if Pashtuns are present to any degree in a district i , regardless of whether they were the majority group, and 0 otherwise. Our main measure relies on Ethnologue. An alternative measure uses GREG.

Population: This is a minimally-modeled gridded population data that incorporates census population data from the 2010 round of censuses. Population estimates are derived by extrapolating the raw census estimates to a series of target years and are provided for the years 2000, 2005, 2010, 2015, and 2020. We use the interpolated data from 2000 till 2015. We then take the logarithm. Source: The Center for International Earth Science Information Network - CIESIN - Columbia University. 2016. Gridded Population of the World, Version 4 (GPWv4): Administrative Unit Center Points with Population Estimates. Palisades, NY: NASA Socioeconomic Data and Applications Center (SEDAC) <http://dx.doi.org/10.7927/H4F47M2C>, last accessed August 28, 2019.

Ruggedness: We calculate the average ruggedness index for every district. While ruggedness refers to the variance in elevation, we also use raw elevation data. Source: Elevation data from NASA Shuttle Radar Topography Mission (SRTM) data set. The data on terrain rugged-

ness is the same that was used in Nunn & Puga (2012), although we use it on a more disaggregated level. The dataset and a detailed documentation are available at <http://diegopuga.org/data/rugged/>, last accessed August 28, 2019.

Significant activities (SIGACTS): These variables measure SIGACTS (Significant Activities) in a given district based on military reports. The SIGACTS version we obtained includes the sum of events per district-year and allows to distinguish events according to the type of (i) combat, (ii) target (forces), and (iii) technology. Regarding (i) and (iii), SIGACTS are defined as direct fire (DF), indirect fire (IDF), and improvised explosive devices (IED). DF attacks can be defined as close combat events that are characterized by the use of weapons like small arms or rocket-propelled grenades. IDF attacks, including mortars and rockets, can be heard within a large area, but are less precise when being launched from great distances. While DF and IDF involve fighters, IEDs involve less risk for the perpetrators. IEDs can be placed around roads and directed against moving targets, for instance pro-government convoys. While DF and IDF are less capital-intensive, IED events capture more capital/technology-based conflict. We can also differentiate between events that are related to casualties and those not related to casualties. Regarding (ii) we can distinguish between actors involved, i.e. whether Afghan or coalition forces have been involved in the event. Source: We use data from Shaver & Wright (2016).

Southern Provinces: Dummy variable which we assign a value of 1 for districts located in one of the two provinces Kandahar and Hilmand, and 0 otherwise.

Sum of Assets (weighted): The number of assets the households possess over a set of assets that is constant over 3 survey waves. This set consists of Radio/Tape, Refrigerator, TV, VCR/DVD, Sewing Machine, Thuraya (any phone), Bicycle, Motorcycle, Tractor/Thresher, and Car. Sum of Assets weighted is the sum of asset weighted by the proportion of households not possessing the specific item. Source: NRVA (CSO, 2005, 2007/08, 2011/12).

Vegetation Health Index (VHI): We compute an index that captures inter-annual variations in drought conditions, the vegetation health index (VHI) of FAO (Van Hoolst *et al.*, 2016). VHI is a composite index joining the Vegetation Condition Index (VCI) and the Temperature Condition Index (TCI, Kogan 1995). Low values of VHI represent drought conditions.

This is a combination of low values of the observed VCI (relatively low vegetation) and higher values of the TCI (relatively warm weather). For details see Van Hoolst *et al.* (2016). The VHI is calculated from data of Advanced Very High Resolution Radiometer (AVHRR) sensors on board of the National Oceanic and Atmospheric Administration (NOAA) and Meteorological Operational Satellite (METOP) satellites. It is superior to simply using precipitation data, which do not directly measure drought conditions, require assumptions about the linearity of the effect and, in particular in Afghanistan, have severe limitations in terms of quality and resolution. The index is based on earth observation data and is available on a monthly basis with a resolution of 1 km^2 . As cultivation and harvest times differ within Afghanistan, we use the yearly average. The remote sensing based index is operationally used to monitor drought conditions in the Global Early Warning System (GEWS), low VHI values indicate drought conditions. For a similar approach to a VHI, see Harari & La Ferrara (2018).

Wheat Price (International): Prices are period averages in nominal US dollars with 2005 as the baseline. We use benchmark prices, representative of the global market. They are determined by the largest exporter of a given commodity. Source: International Monetary Fund (IMF) Primary Commodity Prices database (IMF, 2005-2017).

Wheat Suitability: Seven different soil quality ratings (SQs) are calculated and are combined in a soil unit suitability rating (SR, %). The SR represents the percentage of potential yield expected for a given crop/Land Utilization Type (LUT) with respect to the soil characteristics present in a soil map unit of the HWSD and is depending on input/management level. The FAO-GAEZ (2012) model provides for each crop/LUT a comprehensive soil suitability evaluation for all the soil units contained in the Harmonized World Soil Database (HWSD). This is done by the use of individual soil quality ratings. Source: Global Agro-ecological Zones (GAEZ v3.0) by the Food and Agriculture Organization of the United Nations (FAO-GAEZ 2012). Details are provided on <http://www.fao.org/nr/gaez/about-data-portal/agricultural-suitability-and-potential-yields/en/>, last accessed August 28, 2019. Go to the section “Agro-ecological suitability and productivity” to find the suitability we use and access the data portal for downloads.

B. Descriptive statistics

Table B.1: Descriptive statistics, district level: Part 1

	Observations	Mean	Stand. Dev.	Min.	Max.
(log) BRD	5174	1.11	1.54	0.00	8.20
Small Conflict	5174	0.31	0.46	0.00	1.00
Low Conflict	5174	0.23	0.42	0.00	1.00
Conflict	5174	0.14	0.34	0.00	1.00
War	5174	0.03	0.18	0.00	1.00
(log) Taliban-Civilians BRD	5174	0.08	0.37	0.00	4.14
(log) Taliban-Government BRD	5174	1.05	1.52	0.00	8.20
(log) Government BRD caused by Taliban	5174	0.53	0.94	0.00	8.03
(log) Taliban BRD caused by Government	5174	0.77	1.33	0.00	6.39
With Casualties (SIGACTS)	5174	8.83	36.27	0.00	870.00
Without Casualties (SIGACTS)	5174	27.04	127.68	0.00	5230.00
Direct Fire (SIGACTS)	5174	23.04	124.34	0.00	5181.00
Indirect Fire (SIGACTS)	5174	5.52	19.04	0.00	433.00
Coalition Forces (SIGACTS)	5174	14.34	105.71	0.00	4566.00
Afghan Forces (SIGACTS)	5174	16.44	82.65	0.00	3582.00
IED Share (SIGACTS)	5174	16.03	25.11	0.00	100.00
(log) Wheat Profitability	5174	-0.48	0.46	-2.11	0.01
(log) Opium Profitability - Int. Heroin	5174	-1.52	0.66	-4.61	-0.00
(log) Opium Profitability - Local Opium	5174	-1.04	0.70	-4.61	0.01
(log) Opium Profitability - Int. Complement	5174	-1.30	0.56	-3.17	-0.00
(log) Opium Profitability - Int. Cocaine	5174	-1.15	0.67	-4.61	-0.00
Opium Suitability	5174	0.53	0.18	0.00	1.00
Wheat Suitability	5174	0.55	0.23	0.00	1.00
(log) Cultivation	5174	1.38	2.15	0.00	6.91
(log) Opium Revenue	5149	4.26	5.83	0.00	16.98

Notes: The sample is based on the specification in Table 1, column 1.

Table B.2: Descriptive statistics, district level: Part 2

	Observations	Mean	Stand. Dev.	Min.	Max.
Ruggedness in 1000	5148	299.18	216.54	4.48	877.01
Any Military Base	5174	0.04	0.20	0.00	1.00
Market Access - Opium Market 2D	5174	4.47	1.10	2.24	11.23
Market Access - Opium Market 3D	5174	2.63	0.69	1.33	6.93
Market Access - Luminosity 2D	5174	6.51	4.86	1.85	41.26
Market Access - Luminosity 3D	5174	6.47	4.84	1.85	41.24
Major/Sub Market	5174	0.27	0.44	0.00	1.00
Sum of all Markets	5174	0.40	0.85	0.00	8.00
Any Processing Lab	5174	0.13	0.34	0.00	1.00
Ethnic Trafficking Route	5174	0.52	0.50	0.00	1.00
Distance to Kabul: Linear	5148	277.05	181.54	0.00	817.64
Distance to Kabul - Road 2D	5174	345.03	212.05	0.00	959.78
Distance to Kabul - Road 3D	5174	347.47	213.08	0.00	964.48
Travel Time to Kabul - 2D	5174	7.53	5.91	0.00	28.40
Travel Time to Kabul - 3D	5174	7.57	5.94	0.00	28.45
Pashtuns (Ethnologue)	5174	0.54	0.50	0.00	1.00
Pashtuns (GREG)	5174	0.74	0.44	0.00	1.00
Ethnic Groups - 1 if Mixed	5174	0.59	0.49	0.00	1.00
Ethnic Groups - Number	5174	1.93	0.97	1.00	5.00
Mixed Territory 1996	5174	0.04	0.20	0.00	1.00
Taliban Territory 1996	5174	0.58	0.49	0.00	1.00

Notes: The sample is based on the specification in Table 1, column 1.

Table B.3: Descriptive statistics, household level

	Observations	Mean	Stand. Dev.	Min.	Max.
Dietary Diversity	72224	6.47	1.39	0.00	8.00
Calorie Intake	71634	2625.92	1376.62	1.26	66494.29
Food Insecurity	71634	0.34	0.48	0.00	1.00
Food Expenditures	72643	1927.78	1326.12	0.00	31111.90
Food Exp. Paasche adj.	72643	1893.83	1269.51	0.00	34471.19
Food Exp. Laspeyres adj.	72635	1868.05	1273.41	0.00	31638.37
Sum of Assets	72447	2.17	1.59	0.00	10.00
Sum of Assets Weighted	66621	1.17	1.11	0.00	7.93
Economically Improved	70670	2.81	0.99	1.00	5.00

Notes: The sample is based on the specification in Table 1, column 1.

Table B.4: Type of violence and fighting parties, UCDP GED

Conflict Dyads		
	Frequency	Percent
Government of Afghanistan - Taliban	14,853	95.26
Government of United States of America - al-Qaida	125	0.80
Taliban - Civilians	614	3.94

Type of violence		
	Frequency	Percent
State-based violence	15,084	95.39
One-sided violence	98	0.62
Non-state violence	631	3.99

Notes: Summary of fighting parties and types of violence in Afghanistan provided by UCDP GED between 2002-2014.

Table B.5: Balancing tests - high and low opium suitable districts

	Mean Values per Group		P-Value
	High Suitability	Low Suitability	
Ruggedness in 1000	286.052	342.550	0.000
Distance to Kabul - Linear	248.425	371.647	0.000
Distance to Kabul - Road 2D	311.787	454.068	0.000
Distance to Kabul - Road 3D	314.037	457.126	0.000
Travel Time to Kabul - Road 2D	6.560	10.693	0.000
Travel Time to Kabul - Road 3D	6.597	10.755	0.000
Pashtuns	0.780	0.602	0.000
Mixed Ethnic Groups	0.538	0.742	0.000
Number Ethnic Groups	1.830	2.247	0.000
Mixed Territory 1996	0.030	0.075	0.000
Taliban Territory 1996	0.593	0.527	0.000
Ethnic Trafficking Route	0.557	0.409	0.000
BRD 2000	14.308	11.075	0.172
Luminosity 2000	0.160	0.213	0.322
(log) Population 2000	3.974	2.654	0.000
Wheat Suitability	0.609	0.371	0.000

Notes: Sample based on Table E.1, column 1. To assign a districts to low or high suitability, we use a cut-off of 0.4. In Table F.18 we control for an interaction of all the variables (above the separating line) with a time trend or with time-fixed effects.

C. Geographical overview



Figure C.1: Afghanistan and its neighboring states

Notes: Opium is reported to be mostly trafficked through Iran, Pakistan as well as Turkmenistan according to UNODC.



Figure C.2: Elevation and mountainous terrain in Afghanistan

Notes: The central and north-eastern part of Afghanistan feature the most mountainous terrain. Mountains are correlated with opium suitability, for instance very high altitude areas with a lot of snow are obviously unsuitable, but generally opium can be produced in many places as our map for the suitability indicator shows. We will run regressions with and without the border districts, as well as regressions controlling for elevation or ruggedness (in a flexible way interacted with year dummies) to account for potentially time-varying effects of these factors. Source for elevation data: (U.S. Geological Survey (USGS) Global Multi-resolution Terrain Elevation Data 2010 (GMTED2010), available at <https://lta.cr.usgs.gov/GMTED2010>, accessed 06.04.2018). Source for ADM1 administrative data is <http://www.gadm.org>, last accessed August 28, 2019.



Figure C.3: ADM1 level (provinces) of Afghanistan

Notes: The figure plots the 34 provinces (ADM1 level). Source: Central Statistical Office Afghanistan (<http://afghanag.ucdavis.edu/country-info/about-afghan.html>, last accessed August 28, 2019).

D. Identification using complement prices

In this section, we discuss to which extent we can use additional information on complement prices to learn something about the potential impact of omitted variables. Specifically, we are interested in the biasing effect of an omitted variable OV_{t-1} that varies over time. This bias would affect the opium estimate (b^o) if the omitted variable bias correlates with opium prices and impacts conflict different in low and high suitability districts. Given that the bias would work through affecting the price, and local demand plays a negligible role for overall profits, OV_{t-1} will be any events or shocks that influence opium prices by affecting opium supply.

The underlying idea of using a complement to opium is that a supply shock affects the price of opium and its complement in the opposite direction. Therefore, an omitted variable will also have opposite effects on the estimates regressing the respective prices interacted with suitability on conflict. Section D.1 begins by deriving conditions under which we can exploit this mechanism to learn something about the sign and the bounds of the true underlying relationship between prices and conflict. In Section D.2, we then illustrate the theoretical results using Monte Carlo simulations. Section D.3 finally shows the corresponding proofs and how under stronger assumptions the true parameter is even bounded by the estimates using the opium price and the complement price. We will illustrate everything with regard to opium prices and the respective complements, but the underlying theory and strategy can be applied to any other goods that fulfill the conditions outlined below.

D.1. Theoretical predictions

Set Up: We begin by setting up the price equations:

$$p_{t-1}^o = \eta DS_{t-1} - q_{t-1}^o + \bar{\omega} q_{t-1}^c + \varepsilon_{t-1}^o \quad (1)$$

$$q_{t-1}^o = OV_{t-1} + \xi_{t-1}^o \quad (2)$$

$$p_{t-1}^c = \eta DS_{t-1} - q_{t-1}^c + \bar{\omega} q_{t-1}^o + \varepsilon_{t-1}^c \quad (3)$$

$$q_{t-1}^c = \xi_{t-1}^c \quad (4)$$

Variables p_{t-1}^o and p_{t-1}^c denote the price of respectively opium and the complement at year $t - 1$. Prices are a function of common demand shifters (DS_{t-1}) and supply shocks (OV_{t-1}, ξ_{t-1}). The omitted variable (OV_{t-1}) only affects opium supply so that $\rho(q_{t-1}^c, OV_{t-1}) = 0$. The parameter $\bar{\omega} \in [0, 1]$ indicates the degree to which the two prices react to changes in the supply of the complement. A higher $\bar{\omega}$ indicates that each price reacts stronger to each others supply. To simplify notation, the coefficient of the omitted variable in equation (2) is set to unity. In that case, $\eta > 0$ directly measures the effect of the common demand shifters relative to the omitted variable. Both ξ_{t-1} and ε_{t-1} can be interpreted as (product specific) error terms. Substituting supply in the price equation gives us the following:

$$p_{t-1}^o = \eta DS_{t-1} - OV_{t-1} + \tilde{\varepsilon}_{t-1}^o \quad (5)$$

$$p_{t-1}^c = \eta DS_{t-1} + \bar{\omega} OV_{t-1} + \tilde{\varepsilon}_{t-1}^c \quad (6)$$

with $\tilde{\varepsilon}_{t-1}^o \equiv \varepsilon_{t-1}^o - \xi_{t-1}^o + \bar{\omega} \xi_{t-1}^c$ and $\tilde{\varepsilon}_{t-1}^c \equiv \varepsilon_{t-1}^c - \xi_{t-1}^c + \bar{\omega} \xi_{t-1}^o$. We are interested in estimating the effect of prices on conflict, which we assess using the following regressions:

$$c_{d,t} = \beta(p_{t-1}^o \times s_d) + \gamma(OV_{t-1} \times s_d) + \delta_d + \tau_t + u_{d,t} \quad (7)$$

$$c_{d,t} = b^o(p_{t-1}^o \times s_d) + \delta_d + \tau_t + v_{d,t} \quad (8)$$

$$c_{d,t} = b^c(p_{t-1}^c \times s_d) + \delta_d + \tau_t + v_{d,t} \quad (9)$$

Equation (7) is the true regression and equations (8) and (9) are both two regressions that do not include the unobserved OV_{t-1} . The outcome variable $c_{d,t}$ denotes conflict at year t in district d and s_d the opium suitability in district d . All regressions include district (δ_d) and year (τ_t) fixed effects.

We can use those equations to derive the following estimates:

$$\text{plim } \hat{b}^o = \beta + \tilde{\gamma} \quad (10)$$

$$\text{plim } \hat{b}^c = \theta\beta - \alpha\bar{\omega}\tilde{\gamma} \quad (11)$$

where:

$$\tilde{\gamma} \equiv -\gamma \frac{\sigma_{OV}^2}{\sigma_{p^o}^2}, \quad \theta \equiv \frac{\eta^2 \sigma_{DS}^2 - \bar{\omega}(\sigma_{OV}^2 + \sigma_{\xi^o}^2 + \sigma_{\xi^c}^2)}{\sigma_{p^c}^2}, \quad \alpha \equiv \frac{\sigma_{p^o}^2}{\sigma_{p^c}^2}$$

To make sense of the estimates, we assume $\theta > 0$. Using the definition of θ this restriction is equivalent to:

$$\eta^2 > \bar{\omega} \frac{\sigma_{OV}^2 + \sigma_{\xi^o}^2 + \sigma_{\xi^c}^2}{\sigma_{DS}^2} \quad (12)$$

In other words, we impose that the demand side relative to the supply side must be large enough to discipline θ . As we argue in the main part of the paper, this seems justified for drug prices over the last decades. Since both α and $\bar{\omega}$ are positive, we see that OV_{t-1} affects the estimates in opposite direction. Hence a larger $\tilde{\gamma}$ increases the opium estimate, while it decreases the complement estimate. We will exploit this mechanism in the next section to draw inference on our true estimate β .

Results: In this section we discuss how the asymptotic properties of the estimates behave. More specifically, we derive sufficient conditions to both determine the sign of β and its upper or lower bound. Whenever we refer to $\hat{b}^o$ or $\hat{b}^c$ we refer to its probability limit unless otherwise noted.

First, we can conclude that if we observe both estimates ($\hat{b}^o$ and $\hat{b}^c$) having the same sign, also the true estimate β must have the same sign. More formally:

Corollary 1.

$$\beta < 0 \text{ if } \hat{b}^o, \hat{b}^c < 0 \quad (13)$$

$$\beta > 0 \text{ if } \hat{b}^o, \hat{b}^c > 0 \quad (14)$$

Proof. We prove $\beta < 0$ if $(\hat{b}^o, \hat{b}^c < 0) \wedge (\theta > 0)$ as the second statement follows a similar proof. Since $\hat{b}^o < 0$ from (10) it must be that $(\beta < 0) \vee (\tilde{\gamma} < 0)$. Then if $\tilde{\gamma} < 0$ we must have from (11) that $\theta\beta < 0$ and since $\theta > 0$ it must be that $\beta < 0$. $\square$

This implies that if both the estimates using opium and complements ($\hat{b}^o$ and $\hat{b}^c$) are negative, as we observe in Afghanistan, then also the opium estimate without omitted variable bias $\hat{\beta}$ must be negative.

Second, we can proof that if both $\hat{b}^o$ and $\hat{b}^c$ have the same sign and $\hat{b}^c$ is further away from zero, $\hat{b}^o$ tends to be biased towards zero. More formally:

Corollary 2.

$$\beta < \hat{b}^o < 0 \text{ if } (\hat{b}^c < \hat{b}^o < 0) \quad (15)$$

$$\beta > \hat{b}^o > 0 \text{ if } (\hat{b}^c > \hat{b}^o > 0) \quad (16)$$

Proof. We again prove the first statement as the second follows a similar proof. First, from Corollary 1 we know that $\beta < 0$ if $\hat{b}^c, \hat{b}^o < 0$. Then from (10) we conclude that for $\beta < \hat{b}^c < 0$ it must be true that $\tilde{\gamma} > 0$. Knowing that $\hat{b}^c < \hat{b}^o$ and $\beta < 0$, $\tilde{\gamma} > 0$ is satisfied by construction. $\square$

In other words, the opium estimate $\hat{b}^o$ bounds the true estimate as either an upperbound or a lowerbound if the estimates ($\hat{b}^o, \hat{b}^c$) are respectively negative or positive and the complement estimate $\hat{b}^c$ is higher in absolute value than $\hat{b}^o$.

From Table 1 we know that, apart from being both negative, the complement estimate is more negative than the opium estimate. From Corollary 2 we henceforth learn that our main estimate is biased towards zero in the presence of omitted variable bias. This in turn implies that our estimates are conservative and that the magnitude of the true effect is (potentially) larger.

D.2. Simulation

We now propose a Monte Carlo approach to illustrate how the asymptotic properties regarding sign and bounds of true β behave in the finite sample setup. In the simulation, we vary γ , the

impact of the omitted variable on the outcome variable, in three different ways. First we draw γ from the normal distribution ($\gamma \sim N(0, 1)$) since we do not know the direction of the bias. In the second scenario, we analyze the upward bias, and in the third scenario the downward bias. We simulate a very general data generating process where we assume for the common demand shifters $DS_t \sim N(0, 1)$, and for the omitted variable $OV_t \sim N(0, 1)$. All error terms are also drawn from the standard normal distribution with mean zero and variance equal to one. For the opium suitability, we assume $s_d \sim U[0, 1]$. Finally, we set the impact of the demand shifters on price (η) so that the inequality from Remark 1 holds.

We generate the outcome variable, the conflict in district d at time t using the following

$$c_{dt} = \alpha + \beta(p_{t-1}^o \times s_d) + \gamma(OV_{t-1} \times s_d) + \tau_t + \delta_d + u_{d,t} \quad (17)$$

with $u_{d,t} \sim N(0, 1)$. In each round, we draw 1,000 observations, clustered in 100 districts with 20 time periods, and compute all variables. For each row, we then estimate the following three equations which refer to equations (7 – 9):

$$c_{d,t} = \alpha + \beta(p_{t-1}^o \times s_d) + \gamma(OV_{t-1} \times s_d) + \tau_t + \delta_d + u_{d,t}, \quad (18)$$

$$c_{d,t} = \alpha + b^o(p_{t-1}^o \times s_d) + \tau_t + \delta_d + v_{d,t}, \quad (19)$$

$$c_{d,t} = \alpha + b^c(p_{t-1}^c \times s_d) + \tau_t + \delta_d + v_{d,t}. \quad (20)$$

We set true β equal to zero to analyze the most worrisome case, in which the true effect is insignificant, but due to the omitted variable bias, we may falsely interpret the effect as statistically significant. We want to understand how we can reduce the overrejection of the true null hypothesis that $\beta = 0$ by using information obtained from estimating both b^o and b^c . We are considering three rejection rules:

Rule 1: p-value of $\hat{b}^o$ smaller than 5%.

Rule 2: $\hat{b}^o$ and $\hat{b}^c$ are of the same sign ($\hat{b}^o \cdot \hat{b}^c > 0$).

Rule 3: $\hat{b}^o$ and $\hat{b}^c$ are of the same sign, and the p-values of both $\hat{b}^o$ and $\hat{b}^c$ are smaller than 5%.

The first rule is the naive rule indicating how often we would falsely reject the null hypothesis if we would ignore the presence of omitted variable bias. The second rule captures the idea of upper and lower bounds of the true estimate. The rejection rate under this rule shows the likelihood that true parameter β lies outside bounds of the $\hat{b}^o$ and $\hat{b}^c$ estimates. We showed in Corollary 3 that β must be bounded by $\hat{b}^o$ and $\hat{b}^c$. Since this is an asymptotic result, we want to show how often this holds in a finite sample. In the third rule, we impose additional restriction on the statistical significance of $\hat{b}^o$ and $\hat{b}^c$.

We repeat the simulation 10,000 times and in each round we store the estimates ($\hat{\beta}$, $\hat{b}^o$ and $\hat{b}^c$) and the corresponding standard errors and p-values. We show how likely it is that we would falsely reject the null hypothesis given the results we obtain in a finite sample. We then compare rejection rates obtained by imposing our three rules to the rejection rate based on p-value of the unbiased estimator (p-value of $\hat{\beta}$ smaller than 5%). In this case, the true hypothesis should be rejected in 5% of cases.

Table D.1 shows the mean estimates and rejection rates in the three scenarios. In column 1, γ is drawn from a standard normal distribution. Column 2 shows the upward bias scenario and column 3 the downward bias scenario. While the true estimates $\hat{\beta}$ have mean zero, the average estimates of $\hat{b}^o$ and $\hat{b}^c$ are substantially different from zero in the upward and downward bias scenarios. Moreover, we can see that, on average, $\hat{b}^o$ and $\hat{b}^c$ estimates have opposite signs. In the unknown bias scenario, omitted variable bias differs from zero in every repetition, but has an expected value of zero. Thus, the average estimates $\hat{b}^o$ and $\hat{b}^c$ in column 1 are close to zero.

Panel B in Table D.1 shows the rejection rates for different rejection rules. If we observed the omitted variable we would reject the null hypothesis in 5% of cases. We find that in more than 50% of the cases $\hat{b}^o$ is significantly different from zero. This implies that in at least half of the cases we would incorrectly find a statistically significant impact of opium profitability. The results for rule 2 show that the likelihood of β being outside of the bounds is quite high (35% in the unknown bias scenario, and 25% in both the upwards and downwards bias scenarios). However, if we would impose an additional restriction on both estimates being statistically significant, we would reject the true null hypothesis in roughly 5% of cases. Hence, the probability

Table D.1: Simulation results

	Unknown bias $\gamma \sim N(0, 1)$ (1)	Upward bias $\gamma = -1$ (2)	Downward bias $\gamma = 1$ (3)
<i>A. Average estimates</i>			
$\hat{\beta}$	0.000	-0.000	0.000
$\hat{b}^o$	0.000	0.077	-0.077
$\hat{b}^c$	-0.001	-0.077	0.076
<i>B. Rejection rates</i>			
$p(\hat{\beta} = 0) < 0.05$	0.052	0.055	0.056
Rule 1: significance of b^o			
$p(\hat{b}^o = 0) < 0.05$	0.505	0.709	0.717
Rule 2: same sign of b^o and b^c			
$\hat{b}^o \cdot \hat{b}^c > 0$	0.345	0.250	0.248
Rule 3: same sign and significance of both b^o and b^c			
$p(\hat{b}^o = 0) < 0.05 \wedge p(\hat{b}^c = 0) < 0.05 \wedge \hat{b}^o \cdot \hat{b}^c > 0$	0.052	0.066	0.071

Notes: Simulations with 10,000 repetitions. In each repetition, regressions are based on a dataset with 20 years and 100 districts. $\hat{\beta}$ is the estimate of the impact of opium profitability on conflict obtained from the true regression, i.e. one that controls for the omitted variable (true β equals 0). $\hat{b}^o$ is the estimate of the impact of opium profitability on conflict using the opium price only (without controlling for omitted variable), and $\hat{b}^c$ using the complement price (without controlling for the omitted variable). The first column shows the results when the impact of the omitted variable on conflict (γ) is drawn from a standard normal distribution. The second column shows the results for an upward bias ($\gamma = -1$) and the third column show the results for a downward bias ($\gamma = 1$). Panel A displays the average values of the estimates. Panel B shows the rejection rate of the true hypothesis based on the selected rejection rules.

$p(\cdot)$ indicates that the coefficient estimates are significantly different from zero at the 5%-level.

of rejecting true null hypothesis is more or less equal to the probability of rejecting the true null hypothesis based on unbiased estimates of the true β . We can therefore say with confidence that the true estimate is unlikely to be zero when both the opium estimate and the complement estimates are significant and have the same sign.

We further illustrate simulation results in Figures D.1-D.3. We focus on two scenarios: the graphs on the left side present results for the upward bias, and the graphs on the right side show results for the downward bias. In Figure D.1, we plot the distribution of the estimates. We can clearly see that the estimates based on the opium price and estimates based on the complement price are biased in the opposite directions to each other as stated in Corollary 3.

In Figures D.2-D.3, we show how we improve our inference using complement prices. Given our different rejection rules, the dark red points indicate $\hat{b}^o$ estimates for which we falsely reject the null hypothesis and the light red points the estimates for which we do not reject. Figure D.2 shows the results for rule 1. If we would make our rejection decisions based on p-value of $\hat{b}^o$ only, we would falsely reject the null hypothesis in many cases. Under this rule, the overrejection is largely driven by the bias. For example, most of the false rejections in the upward bias scenario come from estimates that are greater than zero. The rejection of few negative estimates may be attributed to the noise generated by the error terms. Figure D.3 shows that we can substantially reduce the overrejection of the true hypothesis by conditioning on the statistical significance and the sign of $\hat{b}^o$ and $\hat{b}^c$ following Corollary 3. We achieve this by reducing the number of false rejections of estimates driven by the bias. In the upward bias scenario, we substantially reduce the number of false rejections for positive estimates. The rejection rate of negative estimates remains similar as in rule 1, because these rejections are driven by the noise and not the omitted variable bias. Thus, we showed that in the finite sample, the additional information obtained from using complement prices allow us to reduce the occurrence of false rejections of the null hypothesis, which are driven by omitted variable bias. The rejection rate is similar to that obtained by estimating true β , because the remaining rejections are driven by the noise generated by the error terms.

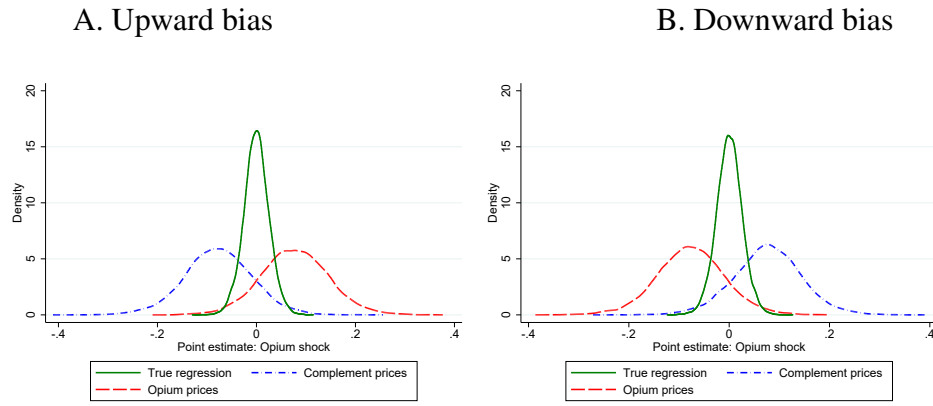
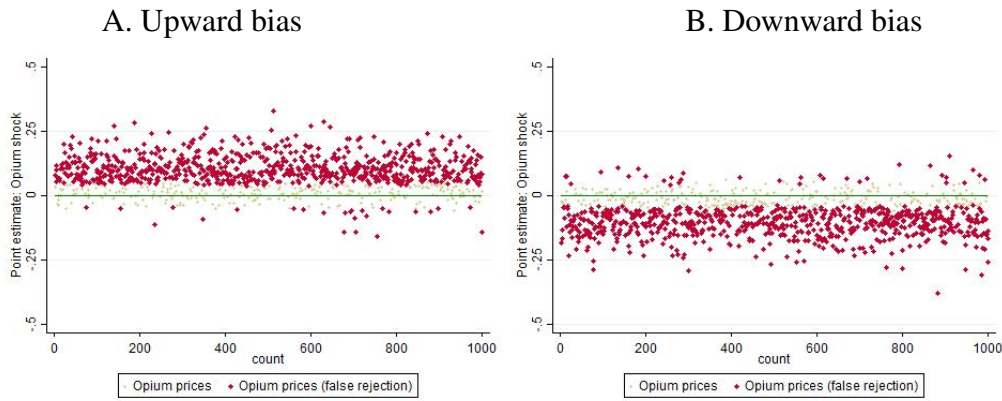
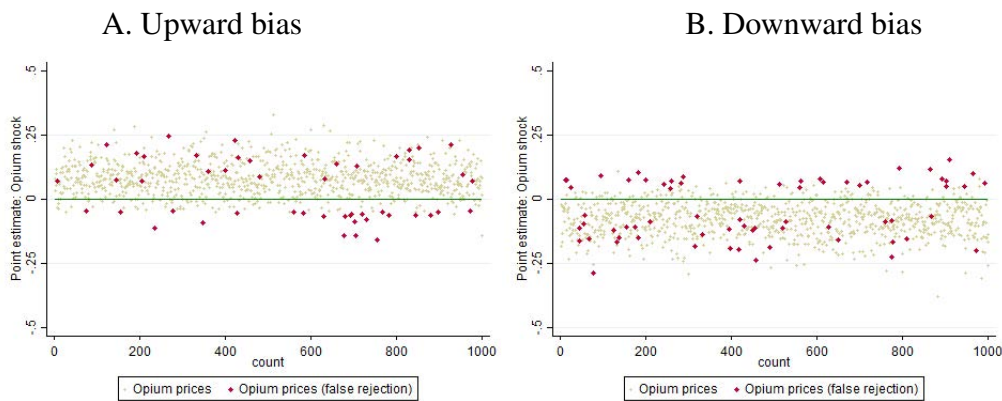


Figure D.1: Distribution of estimates

Figure D.2: Rule 1 - Significance of $\hat{b}^o$ Figure D.3: Rule 3 - Same sign and significance of both $\hat{b}^o$ and $\hat{b}^c$

Notes: Figure D.1 shows the distribution of point estimates (kernel density): $\hat{\beta}$ (green), $\hat{b}^o$ (red), and $\hat{b}^c$ (blue) based on the simulation with 10,000 repetitions. Figures D.2 and D.3 show point estimates: $\hat{b}^o$ that did not lead to the rejection of the null hypothesis (light red), and $\hat{b}^o$ that led to the rejection of the null hypothesis (dark red) based on the simulation with 1,000 repetitions. In each repetition, regressions are based on a dataset with 20 years and 100 districts. The dark red points indicate cases of rejecting true null hypothesis based on different rejection rules. Figure D.2 shows rejections under the following condition: $\hat{b}^o$ is significantly different from 0 at the 5%-level. Figure D.3 shows the rejections under the following condition: $\hat{b}^o$ and $\hat{b}^c$ are of the same sign and they are both significantly different from zero at the 5%-level.

D.3. Derivations and other results

Derivations: We briefly go over the derivations of the estimators stated in (10) and (11). The derivation of $\hat{b}^o$ is akin to the omitted variable bias formula. For the derivation of $\hat{b}^c$ we use both price equations (5) and (6)

$$\hat{b}^o = \frac{\text{cov}(c_{d,t}, p_{t-1}^o \times s_d)}{\text{var}(p_{t-1}^o \times s_d)} \quad (21)$$

$$= \beta \frac{\text{cov}(p_{t-1}^o \times s_d, p_{t-1}^o \times s_d)}{\text{var}(p_{t-1}^o \times s_d)} + \frac{\text{cov}(OV_{t-1} \times s_d, p_{t-1}^o \times s_d)}{\text{var}(p_{t-1}^o \times s_d)} \quad (22)$$

$$= \beta + \frac{\text{cov}(OV_{t-1}, p_{t-1}^o)}{\text{var}(p_{t-1}^o)} \quad (23)$$

$$= \beta + \frac{\text{cov}(OV_{t-1}, p_{t-1}^o)}{\text{var}(OV_{t-1})} \frac{\text{var}(OV_{t-1})}{\text{var}(p_{t-1}^o)} \quad (24)$$

hence,

$$\text{plim } \hat{b}^o = \beta - \gamma \frac{\sigma_{OV}^2}{\sigma_{p^o}^2} \quad (25)$$

$$\hat{b}^c = \frac{\text{cov}(c_{d,t}, p_{t-1}^c \times s_d)}{\text{var}(p_{t-1}^c \times s_d)} \quad (26)$$

$$= \beta \frac{\text{cov}(p_{t-1}^o \times s_d, p_{t-1}^c \times s_d)}{\text{var}(p_{t-1}^c \times s_d)} + \frac{\text{cov}(OV_{t-1} \times s_d, p_{t-1}^c \times s_d)}{\text{var}(p_{t-1}^c \times s_d)} \quad (27)$$

$$= \beta \frac{\text{cov}(\eta DS_{t-1} - OV_{t-1} + \tilde{\varepsilon}_{t-1}^o, \eta DS_{t-1} + \bar{\omega} OV_{t-1} + \tilde{\varepsilon}_{t-1}^c)}{\text{var}(p_{t-1}^c)} \quad (28)$$

$$\begin{aligned} & + \frac{\text{cov}(OV_{t-1}, p_{t-1}^c)}{\text{var}(p_{t-1}^c)} \\ & = \beta \frac{\eta^2 \text{var}(DS_{t-1}) - \bar{\omega} \text{var}(OV_{t-1}) + \text{cov}(\tilde{\varepsilon}_{t-1}^o, \tilde{\varepsilon}_{t-1}^c)}{\text{var}(p_{t-1}^c)} \\ & + \frac{\text{cov}(OV_{t-1}, p_{t-1}^c)}{\text{var}(OV_{t-1})} \frac{\text{var}(OV_{t-1})}{\text{var}(p_{t-1}^c)} \end{aligned} \quad (29)$$

hence,

$$\text{plim } \hat{b}^c = \frac{\eta^2 \sigma_{DS}^2 - \bar{\omega} (\sigma_{OV}^2 + \sigma_{\xi^o}^2 + \sigma_{\xi^c}^2)}{\sigma_{p^c}^2} \beta + \bar{\omega} \gamma \frac{\sigma_{OV}^2}{\sigma_{p^c}^2} \quad (30)$$

Where the explicit variances of the prices are as followed:

$$\sigma_{p^o}^2 = \eta^2 \sigma_{DS}^2 + \sigma_{OV}^2 + \sigma_{\varepsilon^o}^2 + \sigma_{\xi^o}^2 + \bar{\omega}^2 \sigma_{\xi^o}^2 \quad (31)$$

$$\sigma_{p^c}^2 = \eta^2 \sigma_{DS}^2 + \bar{\omega}^2 \sigma_{OV}^2 + \sigma_{\varepsilon^c}^2 + \sigma_{\xi^c}^2 + \bar{\omega}^2 \sigma_{\xi^c}^2 \quad (32)$$

Other results: In this subsection we state the conditions under which β is bounded. We start with a remark of auxiliary value.

Remark 1. If $\hat{b}^o > \hat{b}^c$ then it must be that:

$$\tilde{\gamma} > \frac{(\theta - 1)\beta}{1 + \alpha\bar{\omega}}$$

If $\hat{b}^o < \hat{b}^c$ then it must be that:

$$\tilde{\gamma} < \frac{(\theta - 1)\beta}{1 + \alpha\bar{\omega}}$$

So that:

$$\tilde{\gamma} > 0 \text{ if } (\beta < 0) \wedge (\hat{b}^o > \hat{b}^c)$$

$$\tilde{\gamma} < 0 \text{ if } (\beta > 0) \wedge (\hat{b}^o < \hat{b}^c)$$

We can now state the cases in which β is between $\hat{b}^o$ and $\hat{b}^c$:

Corollary 3.

$$\beta \in (\hat{b}^c, \hat{b}^o) \text{ if } \begin{cases} \hat{b}^c < \hat{b}^o, \beta < 0, \theta > \bar{\theta} \\ \hat{b}^c < \hat{b}^o, \beta > 0, \theta < \bar{\theta}, \tilde{\gamma} > 0 \end{cases}$$

$$\beta \in (\hat{b}^o, \hat{b}^c) \text{ if } \begin{cases} \hat{b}^o < \hat{b}^c, \beta < 0, \theta < \bar{\theta}, \tilde{\gamma} < 0 \\ \hat{b}^o < \hat{b}^c, \beta > 0, \theta > \bar{\theta} \end{cases}$$

where:

$$\bar{\theta} \equiv 1 + \frac{\alpha\bar{\omega}\tilde{\gamma}}{\beta}$$

Proof. As the following proof can be easily extended, we prove the following case:

$$\beta \in (\hat{b}^c, \hat{b}^o) \text{ if } \hat{b}^c < \hat{b}^o, \tilde{\gamma} > 0, \beta < 0, \theta > \bar{\theta}$$

From (10) we observe that $\beta < \hat{b}^o$ if $\tilde{\gamma} > 0$. From (11) we solve $\beta > \hat{b}^c$ for θ and arrive to $\theta > \bar{\theta}$ if $\beta < 0$. Finally Remark 1 shows that in two cases a restriction $\tilde{\gamma}$ follows by construction and is henceforth superfluous. $\square$

This results states that as long as θ is sufficiently diciplined, the true estimate β is between our two estimates $\hat{b}^o, \hat{b}^c$

Finally, we end this subsection by stating the complete result: β having the same sign and being bounded by the two estimates ($\hat{b}^o$ and $\hat{b}^c$). This result is a combination of the sufficient conditions stated in Corrolaries 1 and 3.

Corollary 4.

$$\hat{b}^c < \beta < \hat{b}^o < 0 \text{ if } (\hat{b}^c < \hat{b}^o < 0) \wedge (\theta > \bar{\theta}) \quad (33)$$

$$\hat{b}^o < \beta < \hat{b}^c < 0 \text{ if } (\hat{b}^c < \hat{b}^o < 0) \wedge (\theta < \bar{\theta}) \wedge (\tilde{\gamma} < 0) \quad (34)$$

$$0 < \hat{b}^o < \beta < \hat{b}^c \text{ if } (0 < \hat{b}^c < \hat{b}^o) \wedge (\theta > \bar{\theta}) \quad (35)$$

$$0 < \hat{b}^c < \beta < \hat{b}^o \text{ if } (0 < \hat{b}^c < \hat{b}^o) \wedge (\theta < \bar{\theta}) \wedge (\tilde{\gamma} > 0) \quad (36)$$

where:

$$\bar{\theta} \equiv 1 + \frac{\alpha \bar{\omega} \tilde{\gamma}}{\beta}$$

E. Further results

E.1. Changes in international prices and local revenues

Table E.1: Effect of international price changes on opium revenues, 2002-2014 period

	Outcome: (t) (1)	Outcome: (t) + (t-1) (2)
Opium Profitability (t-1)	2.336 (0.827)	2.489 (0.749)
Number of observations	5149	5085
Adjusted R-Squared	0.482	0.565

Notes: The dependent variable opium revenues is in logarithms. Column (1) presents lagged effects. Column (2) reports lagged and contemporaneous effects by defining the outcome as the moving average, i.e. $(\text{revenues}(t) + \text{revenues}(t-1))/2$. Opium Profitability is defined as the interaction between the normalized drug prices (in logarithms) and the suitability to grow opium. All models include year- and district-fixed effects. Standard errors clustered at the district level are displayed in parentheses.

E.2. Instrumental variable specifications

To put the results and their size into perspective, we also examine the direct effect of district-level opium revenues on conflict. As explained above, computing revenues requires the use of cultivation estimates – partly estimated based on provincial-level data – and regional yields. Thus, these estimations exclude province-times-year-fixed effects that cause weak IV problems, but still include district- and year-fixed effects. While IV estimations can help with measurement error, one potential issue here is that measurement error could also be non-random if, for instance, the data collection was affected by the risk of conflict. Thus, we regard the IV results mostly as a quantification that accompanies our main results.

In addition to providing a simple OLS estimate, we first instrument the endogenous variable opium revenues with our opium profitability measure. We then introduce a second instrument, the interaction of legal opioid prescriptions in the United States with the suitability to grow opium. Prescribed opioids can affect heroin prices in two ways. Heroin prices could increase if a larger legal opioid supply led to a net increase in demand for heroin. The net effect on heroin prices could also be negative if more users substitute heroin with a legal opioid. Empirically, substitution seems to dominate, as a larger supply of legal opioids strongly correlates with

lower heroin prices ($r=-0.83$). Based on this strong correlation, we use legal opioid prescription interacted with the suitability as our second instrument.¹ Technically, while the reduced form approach in Table 1 presents the ITT effect, Table E.2 identifies the local average treatment effect (LATE) for compliers. Having two IVs enables us to compare the two LATEs, as well as estimating an over-identification test when they are jointly included.

Panel A of Table E.2 presents the OLS and IV results, and panel B reports the first-stage estimates. The OLS effect is negative and statistically significant, but rather small in size. The first-stage estimates in columns 2 and 3 reveal that both instruments work well, with their effect being statistically significant at the 1%-level. Both are strong instruments as indicated by the Kleibergen-Paap (Staiger & Stock, 1997) as well as the effective F-statistic (Olea & Pflueger, 2013).² When we include both IVs in one regression, the coefficients keep their sign but the second instrument dominates. This is not surprising, when considering the strong correlation, and given that the change in prescriptions causes part of the change in heroin prices. The Hansen J test in the joint specification does not reject the overidentification restrictions.

Turning to the second-stage results presented in panel A, instrumenting with opium profitability leads to a negative coefficient that is larger than the OLS coefficient and remains statistically significant. The estimate suggests that a 10% revenue increase leads to a decrease in BRD of about 1.5%. Using legal opioid supply instead of heroin prices leads to a very similar, but slightly larger, effect. A 10% increase is associated with a decrease in BRD of about 1.93%. Using both instruments in column 4 leads to a coefficient that is virtually identical with column 3. The LATE using legal opioid supply, which is clearly exogenous to the Afghan conflict, is very similar in size to the effect using the international heroin prices. This also further assures us about the validity of our main approach.³ Applying two different approaches, the reduced form and IV, we consistently find that higher prices are causing a reduction in conflict.

¹ Note that the empirical literature is divided about the general causal relationship between legal opioids and heroin use (Compton *et al.*, 2016).

² Column 4 shows F-statistics of slightly below 10, but the Olea–Pflueger test for weak instruments for a 10% confidence level and a desired weak-instrument threshold τ of 20% rejects the null hypothesis of a weak instrument (Olea & Pflueger, 2013).

³ Figure E.1 shows the partial leverage plot of the corresponding first-stage regression.

Table E.2: OLS, 1st and 2nd stage IV results for Opium Revenue (t-1), 2002-2014 period

	OLS (1)	IV (2)	IV (3)	IV (4)
Panel A: OLS and IV 2nd stage				
(log) Revenue (t-1)	-0.011 (0.005)	-0.153 (0.083)	-0.193 (0.086)	-0.192 (0.086)
Kleibergen-Paap F stat.		16.379	13.047	8.141
Effective F stat.		15.102	12.029	9.992
Hansen J p-val.				0.220
Panel B: 1st Stage				
Opium Profitability (t-1)		2.922 (0.722)		0.149 (0.726)
Legal Opioids (t-1)*Suitability			-15.384 (4.259)	-14.915 (5.581)

Notes: Models include year- and district-fixed effects. The dependent variable is the log of battle-related deaths in (t). Opium profitability is defined as the interaction between the normalized international heroin price (in logarithms) and the suitability to grow opium. Note that the supply of legal opioids is strongly negatively correlated with heroin prices. The number of observations is 5,104 in all columns across both panels. Standard errors are in parentheses (clustered at the district level).

To account for the different timing as shown in Figure 2a, we show the second and first-stage results when we replace revenue in $t - 1$ with the moving average of revenue in $t - 1$ and t in Table E.3 and Table E.4. The two IVs, opium profitability and legal opioid prescription, are again strong as indicated by the F-statistic. Having alternative sources of exogenous variation also enables us to compare the LATE of the different IVs. The local effects do not differ much either in terms of magnitude or statistical significance.

Table E.3: IVs for revenue in (t)+(t-1), 2002-2014

	(log) BRD (1)	1 if ≥ 5 (2)	1 if ≥ 10 (3)	1 if ≥ 25 (4)	1 if ≥ 100 (5)
Panel A: Opium Price Shock (t-1) as IV					
(log) Revenue:(t)+(t-1)	-0.173 (0.099)	-0.049 (0.030)	-0.045 (0.030)	-0.020 (0.022)	-0.004 (0.010)
Kleibergen-Paap F stat.	11.045	11.045	11.045	11.045	11.045
Panel B: Legal Opioids (t-1) as IV					
(log) Revenue:(t)+(t-1)	-0.252 (0.133)	-0.075 (0.039)	-0.060 (0.036)	-0.032 (0.025)	-0.011 (0.011)
Kleibergen-Paap F stat.	6.671	6.671	6.671	6.671	6.671

Notes: Models include year- and district-fixed effects. The dependent variable is conflict in (t) operationalized as indicated in the column heading. Opium revenues is operationalized as the moving average between (t) and (t-1). The number of observations is 5,085 across columns and panels. Standard errors are in parentheses (clustered at the district level).

Table E.4: Corresponding 1st stage results for revenues (t)+(t-1), 2002-2014

	Opium Profitability (1)	Legal Opioids (2)
Opium Profitability (t-1)	2.489 (0.749)	
Legal Opioids (t-1)		-11.489 (4.448)

Notes: Models include year- and district-fixed effects. The dependent variable is opium revenue or cultivation as indicated in the panel heading. The corresponding IVs are indicated in the column heading. The number of observations is 5,085 across columns. The number of observations is 5,174 across columns and panels. Standard errors are in parentheses (clustered at the district level).

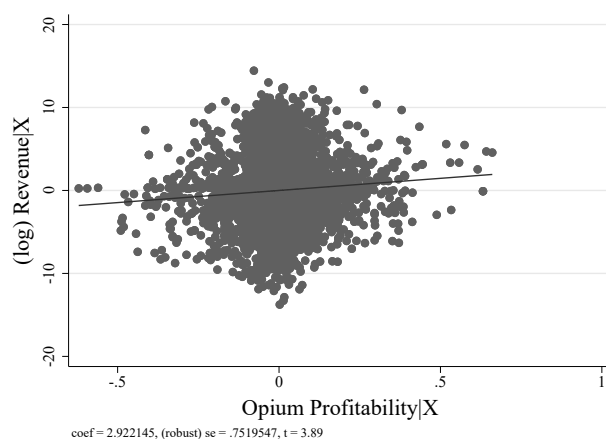


Figure E.1: 1st stage IV results for (log) Revenue

E.3. Aggregate results

Our results do not rule out that local Taliban forces use part of the revenue extracted from the opium business to finance anti-government conflict and attacks. Local revenues could partly be used for violent operations if there are relevant targets within a district. Of course, revenues need not fully remain within the district, and could be pooled to enable countrywide operations. Figure E.2 does not indicate such a mechanism at large scale. On average, an increase in opium revenue correlates with a decrease in casualties.

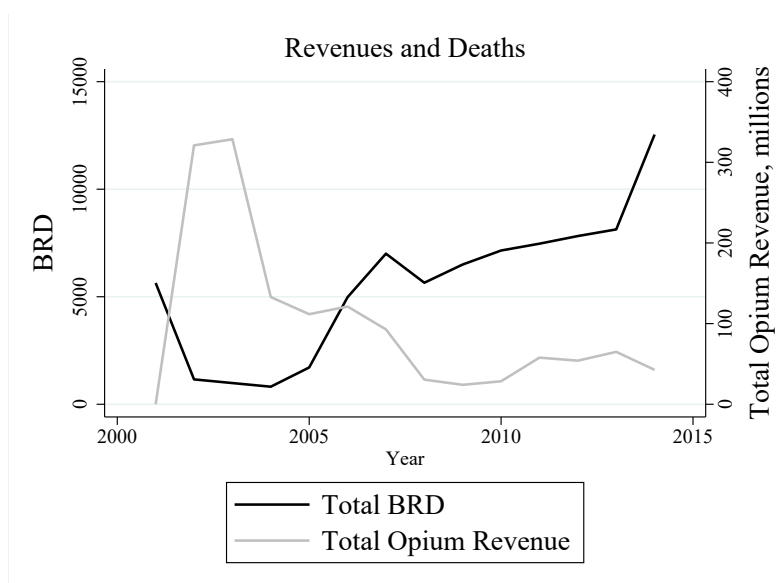


Figure E.2: Correlation between opium revenues and battle-related deaths at country level

Tables E.5 and E.6 show regressions aggregating all our data at the provincial level. Again, we find positive coefficients of opium profitability on opium revenues and a negative coefficient for opium profitability on conflict measured by the logarithm of battle-related deaths. We would expect the opposite pattern if higher revenues in one district were causing more conflict in neighboring districts of other provinces.

Table E.5: Effect of income shocks on opium revenues, at the province level 2002-2014

	Outcome: (t) (1)	Outcome: (t) and (t-1) (2)
Opium Profitability (t-1)	5.885 (3.203)	5.461 (3.078)
Adjusted R-Squared	0.608	0.678

Notes: Models include province- and year-fixed effects. The dependent variable opium revenues in (t) is in logarithms. The number of observations is 442 across columns. Standard errors clustered at the province level are displayed in parentheses.

Table E.6: Additional results using normalized drug prices, at the province level 2002-2014

	Local Opium Price (1)	International Heroin Price (2)	Complement Price (3)
Opium Profitability (t-1)	-0.290 (0.299)	-0.717 (0.566)	-1.101 (0.648)
Adjusted R-Squared	0.722	0.723	0.726

Notes: Models include province- and year-fixed effects. The dependent variable is the log of BRD in (t). Opium Profitability is defined as the interaction between the normalized drug prices (in logarithms) as indicated in the column heading and the suitability to grow opium. The number of observations is 442 across columns. Standard errors are in parentheses (clustered at the province level).

E.4. Counter-factual when freezing prices

In Figure E.3, we fix price levels at the year 2001 to compute a hypothetical counterfactual if prices would not have fallen as much as they did over our sample period. To do so, we use the β estimate from the baseline specification in Table 1, panel B, column 1 on conflict intensity. More precisely, for a given t we calculate the following:

$\hat{BRD}_{t,2001} = \sum_d BRD_{d,t} + \beta \times suitability_d \times BRD_{d,t} \times \frac{price_{2001} - price_{t-1}}{price_{t-1}}$. The figure illustrates that without the strong decline in heroin prices, there would have been an increase in battle-related-deaths, but to a smaller extent than the drastic rise that the country actually experienced.

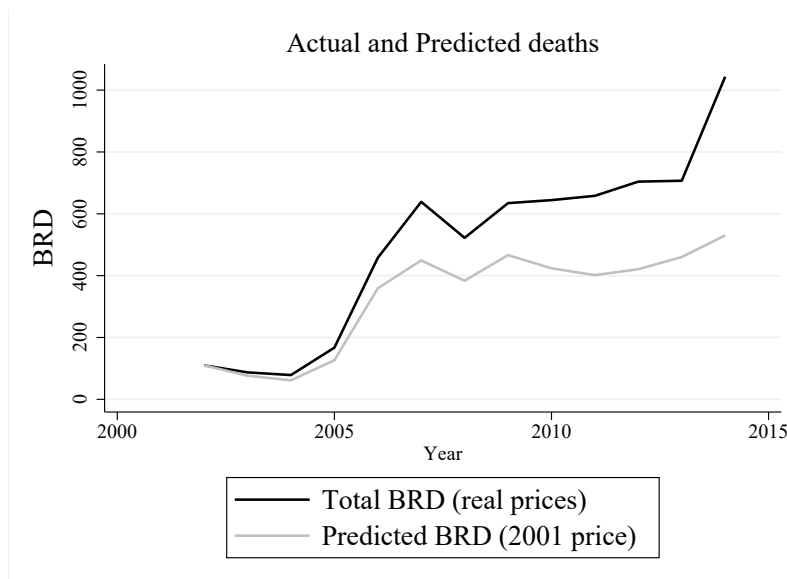


Figure E.3: Counterfactual number of battle-related deaths when fixing prices at higher initial level

E.5. Opportunity costs and shifts in conflict tactics

Increased opportunity costs due to better outside options could not only affect the intensity and frequency of violence but also influence the types of violent activities the actors engage in. Shapiro (2012; 2013) argues that terrorist decision-making can be analyzed as that of rational agents. They try to achieve their goals as best as possible with the limited means they have and adjust their plans to economic shocks such as better outside opportunities (Benmelech *et al.*, 2012). There are several possible explanations why insurgents would also rationally adjust their tactics to changes in opium profitability. Better outside options increase opportunity costs of fighting and reduce the labor supply that local Taliban leaders can draw on for their insurgent activities. In addition, local Taliban leaders are themselves involved in opium production and trade and might thus adapt their tactics to avoid hurting these profitable side activities.

We can derive the following predictions regarding combat tactics. First, rational insurgent leaders are likely to shift from more to less labor-intensive types of combat (as argued in, e.g., Wright, 2018; Bueno de Mesquita, 2013). The scarcer labor becomes, the more planning will shift towards activities with a lower expected death toll. In the SIGACTS data, direct fire (DF)

refers to open combat events with direct enemy contact, thus linked to a high likelihood of own casualties. Indirect fire (IDF), in contrast, comprises rocket and mortar attacks perpetrated at long-range with a lower likelihood of own losses (see, e.g., Blair & Wright, 2021 for detailed definitions). The dataset also allows us to distinguish whether an event is actually linked to casualties. We take this as a second, imperfect proxy for the a priori likelihood of (labor) losses when the insurgent leaders planned the event. Figure E.4, panel A, reports results for events classified according to their labor intensity, following our baseline regression equation with international heroin prices. As outside opportunities increase, direct fire and events with casualties decline more than the less labor-intensive types of combat.

Insurgents might also change their tactics concerning the target type that is attacked (see, e.g., Vanden Eynde, 2018; Wright, 2018). Trebbi *et al.* (2017) show for the Afghan conflict that there is considerable heterogeneity in the prevalence of target types over time. We can think of soft targets as those with a high likelihood of success and hard targets as those where considerable resistance or defeat is more likely. With less labor and the possibility to profit from opium production, local insurgent leaders might be more hesitant to attack hard targets like military bases or foreign military forces. Instead, they might focus more on softer targets like Afghan forces or civilians. Using data from SIGACTS and UCDP GED, we distinguish those target types. Figure E.4, panel B, shows that the smallest and insignificant decline is in attacks on civilians, whereas there seems to be no significant difference between combats with the coalition or Afghan forces.

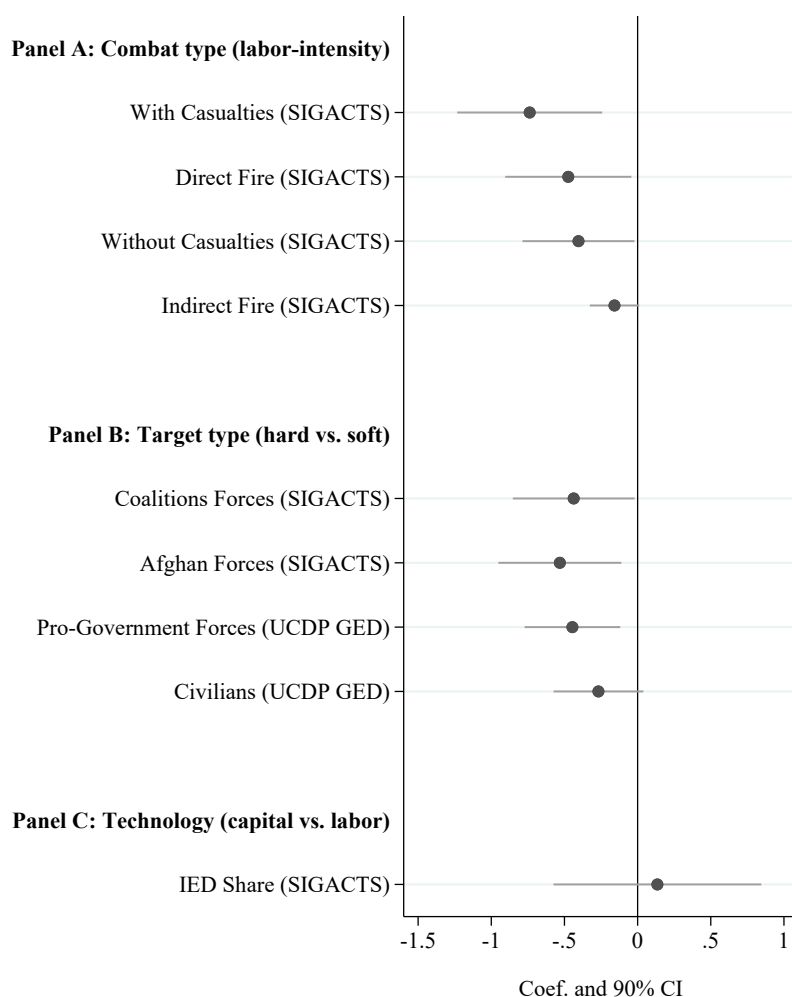


Figure E.4: Distinction by type of combat, target and technology

Notes: The figure plots coefficient estimates (along with 90% confidence intervals) of opium profitability for nine separate regressions with outcomes as indicated on the left-hand side. The outcome variables are based on SIGACTS and UCDP GED, details on their construction can be found in Appendix A. Outcomes in panels A and B are standardized to have a mean value of zero and a standard deviation of one. For reporting, IED share is divided by 10. All regressions follow the baseline regression equation with international heroin prices in Table 1, panel B. Corresponding regression results are in Table E.7 for panel A, in Table E.8 for panel B, and in Table E.9 for panel C.

Finally, with fewer fighters available, insurgents could switch to more technology-intensive attacks, as suggested by Blair & Wright (2021) and Vanden Eynde *et al.* (2017). Combat events using improvised explosive devices (IED) are distinct in requiring both relatively more capital and relatively more skilled labor. This leads to two opposing predictions. On the one hand, higher prices for labor, in general, would predict a shift towards those more capital-intensive IED attacks. On the other hand, the technical specialists needed for IEDs might represent

the type of skilled labor profiting the most from the better outside opportunities, leading to a relative decrease in IED attacks. Figure E.4, panel C, shows that overall, there is a positive coefficient for the share of IEDs, but it is small and not statistically significant.

For all types of events in panels A and B, irrespective of how we classify them, we find coefficient estimates to be negative as opium profitability increases. This is also the case if we consider the total of all significant activities as presented in Table E.7, column 1. There is some interesting evidence for tactical shifts, most clearly for shifts from more to less labor-intensive combat types. This is broadly in line with evidence by Sonin & Wright (2023) that better resource endowments of rebel groups are related to the use of more technology. The results further highlight the importance of opportunity cost effects in shaping the conflict. The next section continues to explore the importance of opportunity costs relative to contest effects.

Table E.7: Distinction by combat-type: Labor intensity, SIGACTs, 2002-2014

Labor intensity:	High		Low	
	All Events	With Casualties	Without Casualties	IDF
	(1)	(2)	(4)	(5)
Opium Profitability (t-1)	-0.494 (0.251)	-0.737 (0.301)	-0.474 (0.262)	-0.404 (0.233)
Adjusted R-Squared	0.493	0.534	0.442	0.469

Notes: Models include province-times-year- and district-fixed effects. The dependent variable is a specific type of conflict activity as indicated in the column heading. Outcomes are standardized to have a mean value of zero and a standard deviation of one. The number of observations is 5,174 across columns. Standard errors are in parentheses (clustered at the district level).

Table E.8: Distinction by target-type: Soft vs. hard, UCDP GED and SIGACTs, 2002-2014

	Coalition	Forces Afghan	Pro-Gov.	Civilians
	(1)	(2)	(3)	(4)
Opium Profitability (t-1)	-0.437 (0.253)	-0.531 (0.255)	-0.446 (0.199)	-0.268 (0.186)
Adjusted R-Squared	0.372	0.386	0.658	0.200

Notes: Models include province-times-year- and district-fixed effects. The dependent variable is a specific type of conflict as indicated in the column heading. Outcomes are standardized to have a mean value of zero and a standard deviation of one. The number of observations is 5,174 across columns. Standard errors are in parentheses (clustered at the district level).

Table E.9: Distinction by technology: Share of DFs and IEDs, SIGACTs, 2002-2014

	DF Share (1)	IED Share (2)
Opium Profitability (t-1)	-1.983 (6.414)	0.135 (0.432)
Adjusted R-Squared	0.381	0.275

Notes: Models include province-times-year- and district-fixed effects. The dependent variable is the share of DFs or IEDs of all SIGACTs events as indicated in the column heading. The number of observations is 5,174 across columns. Standard errors are in parentheses (clustered at the district level).

Table E.10: Types of fighting, UCDP GED, 2002-2014

Conflict Actor:	All	Taliban vs. Civil.	Taliban vs. Gov.		
BRD:	Any	Talib.&Civil.	Talib.&Gov.	Taliban	Government
	(1)	(2)	(3)	(4)	(5)
Opium Profitability (t-1)	-0.675 (0.296)	-0.098 (0.069)	-0.677 (0.302)	-0.539 (0.187)	-0.521 (0.274)
Adjusted R-Squared	0.649	0.200	0.658	0.555	0.596

Notes: Models include province-times-year- and district-fixed effects. The dependent variable is the log of BRD in (t) for a specific type of conflict operationalized as indicated in the column heading. The number of observations is 5,174 across columns. Standard errors are in parentheses (clustered at the district level).

F. Sensitivity analysis

Timing of shocks: Referring to Figure 2a, price changes in $(t-1)$ are most likely affecting cultivation decisions in $\text{summer}(t-1)$, $\text{winter}(t-1)/(t)$, and $\text{summer}(t)$, and thus also labor demand and revenues in (t) . Caulkins *et al.* (2010, p. 9) suggest that “the largest driver of changes in hectares under poppy cultivation is not eradication or enforcement risk, but rather last year’s opium prices.” Taking contemporaneous prices in (t) is problematic with yearly price and conflict data. Using the price in (t) would introduce reverse causality, as price changes later in the year can be affected by earlier conflict. Moreover, we need to allow for time for price shocks to affect local production. We consider different lag structures by including opium profitability in periods $t+1$, t , and $t-1$ at the same time (either all three or in groups by two) in Table F.1, with

$t + 1$ testing for pre-trends. Table F.1 shows that opium profitability in t and $t - 1$ is related to a conflict-reducing effect, while the lead effect of international opium prices in $t+1$ interacted with the suitability to grow opium has no significant effect on conflict. This is reassuring and supports the causal order and mechanism that we hypothesize. Table F.2 shows that including the contemporaneous and lagged variables individually yields very similar coefficients, with slightly larger coefficients for our preferred timing $t - 1$.

Outcome variable (conflict onset and ending): We consider heterogeneous effects of opium profitability on onset and ending of conflict events. Bluhm *et al.* (2021) point to the importance of differentiating between the probability of switching from one conflict state to another as, for instance, from peace to conflict versus from conflict to peace.⁴ Thus, we also measure the effects for conflict incidence (panel A), onset (panel B), and ending (panel C) in separate models. Results are presented in Table F.3. Panel A verifies our main finding with a linear probability model by showing similar results when using conditional logit. In panel B, we find that opium profitability consistently reduces the likelihood of a conflict onset for conflict measured up to a threshold of 25 battle-related deaths. For conflict ending, we only find a significantly positive effect for smaller conflicts. These results indicate that a positive income shock and more opium cultivation raise the likelihood that an ongoing small conflict ends, and reduces the likelihood that conflicts break out.

Ethnic fractionalization: Table F.4 examines whether the conflict-reducing effect is stronger or weaker depending on how many ethnic groups are present. We compute both a binary indicator for whether one or more groups are present, as well as a continuous variable counting the number of groups in a district. This is done based on ethnic homelands from Weidmann *et al.* (2010), as well as using survey based ethnicity measures. All interactions turn out to be insignificant, supporting the notion that after 2001 conflicts occurred mainly between pro-government and pro-Taliban factions, not between ethnicities.

⁴ Berman & Couttenier (2015), for instance, argue that conflict persistence is very low at their level of analysis (a cell equivalent to 55 times 55 km at the equator) compared to country level data. Consequently, they do not include the lagged dependent and rather estimate separate models for onset and ending.

Modifications of the treatment variable: We use multiple modifications of our treatment variable, both by replacing drug prices and crop suitabilities with alternative measures. Tables F.5 and F.6 are equivalent to our main results presented in Table 1 apart from the fact that drug prices are not normalized in Table F.5 and in Table F.6 the prices are not in logarithms. Table F.7 shows the results with de-trended opium prices, which exhibit less variation, but support the main finding. In Table F.8, we use the deviation of the international prices from their long-term mean.⁵ This is a first attempt to rule out that our results are driven by the long-term negative trend in international drug prices as visible in Figure 2b. We find our results to be unaffected by all these choices. With regard to the suitability, we replace the population-weighted suitability for opium and wheat with an unweighted version (see Table F.10). Weighting is important as population density differs strongly across Afghanistan, but causes potential bad control problems due to endogenous migration. While the wheat profitability turns insignificant, the results for opium profitability remain unaffected for all specifications.

Finally, we dichotomize the levels of the interaction. This reduces the complexity of the DiD-like interpretation. In panel A of Table F.11 we dichotomize the suitability based on the sample median. This allows us to interpret a price increase for two groups of districts, i.e., suitable (above the median) and less suitable (below the median). In panel B both variables are dichotomized based on the respective sample median. The coefficient in panel A indicates that a 10% increase in prices leads to about a 2.3% decrease in battle-related deaths in districts with a high suitability. Panel B finds that changing from a low- to a high-price-period reduces deaths by about 50% in districts with high suitability. Across all columns, the results are robust to this adaption.

Empirical model: First, we show our main results with a less restrictive set of fixed effects in Table F.12 for the different prices in panels A to D. The results using only district- and year-fixed effects all point in the same direction, with somewhat smaller coefficients. Again, all four prices consistently indicate a negative effect, both when looking at conflict intensity and incidence. The fact that the effects in our baseline specification become larger in absolute

⁵ Specifically, we use the mean over the entire observation period. Due to data restrictions we cannot calculate the mean over a longer term.

terms, when using province-times-year-fixed effects, suggests that the fixed effects succeed in eliminating biasing variation. In Table F.13 we test for population-density-specific heterogeneity of the main effect. The table shows that using the unweighted opium profitability measure interacted with the population density yields a negative coefficient, in line with the expectation that suitability matters more with higher density. We also run a placebo test by interacting opium profitability (unweighted) with a dummy variable taking value one for districts with a very low population density (below the 10th or 20th percentile). The interaction terms turn out to be positive, indicating that the negative effect is weaker when population density is very low. Note that in both types of tests, the interaction terms are far from being statistically significant.

Outcome (cultivation and revenue): In Sections 4 and 5, we show that there is a strong effect of prices on opium revenues. Table F.14 also supports the positive effect of a higher opium profitability on opium cultivation (in hectares), with positive coefficients that are marginally insignificant in column 1 and significant at the 5%-level in column 2 when considering both periods that are most likely affected by the price change. This is not surprising as opium revenues are affected through changes in price and quantity produced, and cultivation only by the latter.

Standard errors: In a next step, we use different choices on how to cluster standard errors. In the baseline models we used the district level, allowing for serial correlation over time within a district. In Table F.15 we use two-way clustering, i.e., district and year clusters in panel A and province and year clusters in panel B (Cameron *et al.*, 2011). Clustering at the province level is problematic as the number of clusters might be too small, which can lead to the over- or under-rejection of the null hypothesis (Cameron & Miller, 2015). Instead, we use the wild-cluster bootstrap method with the null imposed with 1000 replications and Webb's weights (Webb, 2013), which has been shown to provide valid inference even for few clusters. Figure F.3 plots the distribution of the bootstrap estimates. The null hypothesis of no effect is rejected both when using the international heroin price or the complement price index at least at the 5%-level.

Covariates and trends: Our specifications so far only include wheat profitability and different fixed effects as covariates. It is natural to first compare these results to results without those main covariates. In Table F.16, we find our results to remain robust when we exclude wheat profitability, with slightly more negative coefficients. Table F.16 also shows the tables including the coefficient of wheat profitability for comparison. To account for the persistence of conflict, we include the lagged dependent variable in a next step. Opium profitability remains negative in all columns and statistically significant in columns 1 to 3 (see Table F.17). In Table F.18 (panel A) we add a baseline set of pre-determined covariates such as luminosity and population as well as an exogenous measure of droughts, the VHI. In further specifications (panel B), we also allow for time-varying effects of time-invariant district-specific control variables.⁶ One concern with our specification is that the time trends in prices interact not only with opium suitability, but also with other district characteristics. One way to model this is by adding interactions between these characteristics and a time trend. Another more flexible way is to interact the time invariant control variables with year dummies (panel C). This last specification allows for fully flexible trends interacting with a wide range of district features. The coefficients of our treatment variable are remarkably stable, ranging from -0.675 in the baseline (column 1, panel B, Table 1) to -0.694 in the most flexible specification for conflict intensity. They also remain significant with p-values below at least 0.1 for all conflict proxies (with the sole exception of the category “war”).

Sensitivity to outliers: In Table F.20 we drop potential outliers. In panel A we exclude all border districts from the specification as they could be either very different to other districts or shocks in neighboring countries could affect border districts in a different fashion. For instance, we expect a large share of trafficking to occur close to the border. This could drive the results if international price increases would not reach the average farmer but only the traders, which are closer to the final customer along the supply chain. We find that our results are not driven by this particular group of districts. In panel B, we drop the two southern provinces Kandahar and Helmand and find our results to remain robust to this choice. These provinces are of specific

⁶ The set of time-invariant covariates includes Ruggedness, Ethnic Trafficking Route, Pashtuns, Mixed Ethnic Groups, Taliban Territory 1996, Mixed Territory 1996, Distance Linear, Distance 2D and 3D, Travel Time 2D and 3D (all distances to Kabul).

interest for a number of reasons. First, the Taliban had their origins in the southern region and are thus likely to still have a strong support base here. Second, these provinces are known to be the largest producers of opium. Third, because of their direct connection to Pakistan, which is not only important in relation to trafficking routes but is also a major base of military support for the Taliban.

Apart from these rather obvious heterogeneous groups, we systematically investigate whether results are driven by a particular province or year. Figure F.4 reports the coefficients and the 90% confidence intervals when we drop each year or province one at a time. All coefficients remain stable and within a narrow band.

Randomization: One of the important points raised by Barrett & Christian (2017) is that non-linear trends in the time series of Bartik/shift-share like instruments can be problematic. We address this by looking at prices of different drugs and different versions of these prices (de-trended, log vs. non-log). To further rule out that the results are driven by non-linear trends, we implement further randomization placebo tests. We first randomize the time-varying variable (international heroin price) across years, and in a second specification randomize the district-specific suitability across districts. We would be worried if any of these specifications would create a negative effect similar in magnitude to our treatment effect. Figure F.5 plots the distribution of the coefficients generated by 5'000 randomizations per test along with the actual coefficient. We can also use this to conduct a randomization inference (RI) exercise, in which we compare how many of the random draws generate coefficients that are more negative than ours in order to compute an RI p-value. Reassuringly, we find that if the treatment was randomized according to the two different strategies, the simulated coefficients are always centered around zero. The p-values computed using two-sided symmetric randomization inference are 0.021.

Changes in US military strategy that increases reliance on opium revenues: We validate the importance of opportunity costs by using an important policy change in the foreign coalition's military strategy. This helps us to verify the importance of the opium economy in providing jobs. More importantly, this sheds some light on the effectiveness of nation building

efforts and foreign military interventions, linking our study to the literature on nation building, as for instance Berman *et al.* (2011a) for Iraq and Dell & Querubin (2018) for Vietnam. These studies often consider a distinction between strategies focusing on the use of firepower and military force, and strategies based on winning “hearts and minds” by investing money and providing services and public goods like security. Obviously, each conflict is different, but nonetheless studying the successes and failures often can provide important lessons for the future and other contexts.

In Afghanistan, the coalition forces initially provided strong financial support to existing warlords and local strongholds from roughly 2001 to 2005 to build a strong anti-Taliban coalition. Rough estimates speak of several “hundred thousands of men” being armed as part of local militias, and more than 60% of provincial governors being “leaders of armed groups and most of the remaining ones had links to the latter” (Giustozzi, 2009, p. 91). Around 2005, the coalition switched their strategy towards a nation-building approach that attempted to pacify and “clean” Afghan politics. In this process, intense pressure on the Afghan government forced political leaders and governors to abandon their connection, as well as their support for militias, causing many trained and armed men to lose their main source of income (Giustozzi, 2009, p. 94 ff.). This change in strategy also coincides with the resurgence of the Taliban.

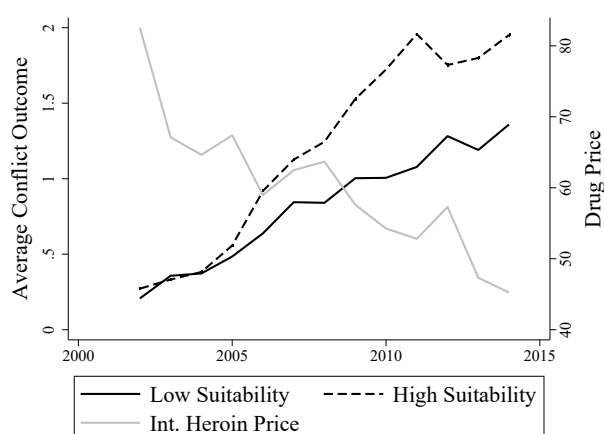


Figure F.1: Variation of conflict across high and low suitability districts over time

Notes: To assign a districts to low or high suitability, we use a cut-off of 0.4. See Appendix F, Figure F.2 for an alternative cut-off of 0.3. Inferences do not depend on this choice.

There is an analogy to the order of events in Iraq, where the de-Baathification process dissolved the Iraqi army and stopped all senior and mid-level party officials from joining the

new army and security services. Various experts assess that this “drove many of the suddenly out-of-work Sunni warriors into alliances with a Sunni/anti-American insurgency” that later joined forces like ISIS. They speak of the “pervasive role played by members of Iraq’s former Baathist army” and estimate that “25 of ISIS’s top 40 leaders once served in the Iraqi military.”⁷ Figure F.1 shows the correlation between prices and conflict in low and high suitability areas. It is clearly visible that, around the approximate timing of this change, the relationship between drug profitability and conflict becomes much stronger in high suitability districts. Dissolving the militias eliminates many reasonably paid jobs, which increases the reliance on income from the opium economy. Hence, the results contrast Berman *et al.* (2011b), who use survey results for two years and find no effect of unemployment. It provides further evidence for the importance of the opportunity cost mechanism in Afghanistan. This also highlights an important trade-off between “cleaning” the state and non-state armed groups as well as fighting the production of an illegal resource at the same time.

Robustness for Figure F.1

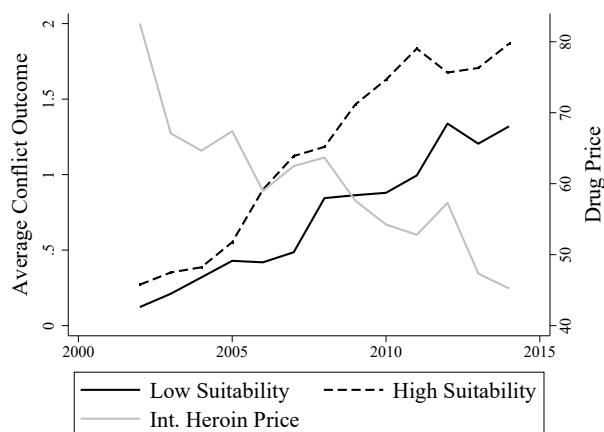


Figure F.2: Variation of conflict across high and low suitable districts over time

Notes: To assign a district to low or high suitability, this figure use an alternative cut-off of 0.3.

⁷ See <http://time.com/3900753/isis-iraq-syria-army-united-states-military/>, <https://www.reuters.com/investigates/special-report/mideast-crisis-iraq-islamicstate/>, <https://www.independent.co.uk/news/world/middle-east/how-saddam-husseins-former-military-officers-and-spies-are-controlling-isis-10156610.html> and <http://nationalpost.com/news/world/how-the-catastrophic-american-decision-to-disband-saddams-military-helped-fuel-the-rise-of-isis/>, last accessed August 28, 2019. A detailed report about “Lessons of De-Baathification in Iraq” is by Sissons and Al-Saiedi, available at <https://www.ictj.org/publication/bitter-legacy-lessons-de-baathification-iraq>, last accessed August 28, 2019.

Different timing of the shocks

Table F.1: Leads and lags, 2002-2014

	Not including: Wheat Profitability (1)	Including Wheat Profitability (2)
Panel A: Including t-1, t, and t+1		
Opium Profitability (t+1)	-0.066 (0.251)	-0.045 (0.252)
Opium Profitability (t)	-0.660 (0.320)	-0.609 (0.321)
Opium Profitability (t-1)	-0.773 (0.289)	-0.459 (0.329)
Number of observations	4776	4776
Adjusted R-squared	0.648	0.649
Panel B: Including t and t+1		
Opium Profitability (t+1)	-0.122 (0.254)	-0.116 (0.251)
Opium Profitability (t)	-1.141 (0.311)	-0.875 (0.324)
Number of observations	4776	4776
Adjusted R-squared	0.648	0.649
Panel C: Including t-1 and t		
Opium Profitability (t)	-0.273 (0.249)	-0.266 (0.249)
Opium Profitability (t-1)	-0.668 (0.286)	-0.405 (0.303)
Number of observations	5174	5174
Adjusted R-squared	0.649	0.649

Notes: Models include province-times-year- and district-fixed effects. The dependent variable is the logarithm of BRD in (t). Wheat profitabillity includes both leads and lags. Standard errors are in parentheses (clustered at the district level).

Table F.2: Timing of shocks, 2002-2014

	(log) BRD (1)	1 if ≥ 5 (2)	1 if ≥ 10 (3)	1 if ≥ 25 (4)	1 if ≥ 100 (5)
Panel A: Contemporaneous effect					
Opium Profitability (t)	-0.608 (0.246)	-0.168 (0.075)	-0.161 (0.074)	-0.130 (0.066)	-0.021 (0.033)
Adjusted R-squared	0.649	0.501	0.483	0.454	0.309
Panel B: Lagged effect					
Opium Profitability (t-1)	-0.675 (0.296)	-0.167 (0.090)	-0.191 (0.085)	-0.147 (0.075)	-0.040 (0.037)
Adjusted R-squared	0.649	0.501	0.484	0.453	0.310

Notes: Models include province-times-year- and district-fixed effects. The dependent variable is conflict in (t) operationalized as indicated in the column heading. The number of observations is 5,174 across columns. Standard errors are in parentheses (clustered at the district level).

Outcome variable (onset and ending)

Table F.3: Conditional logit - incidence, onset and ending, 2002-2014

	1 if ≥ 5 (1)	1 if ≥ 10 (2)	1 if ≥ 25 (3)	1 if ≥ 100 (4)
Panel A: Incidence				
Opium Profitability (t-1)	-1.316 (0.606)	-1.751 (0.673)	-1.631 (0.905)	-1.490 (1.462)
Number of observations	4004	3523	2444	806
Pseudo R-Squared	0.270	0.271	0.272	0.215
Panel B: Onset				
Opium Profitability (t-1)	-0.795 (0.598)	-1.731 (0.717)	-1.130 (0.905)	-0.670 (1.476)
Number of observations	2960	2750	2006	714
Pseudo R-Squared	0.133	0.136	0.148	0.149
Panel C: Ending				
Opium Profitability (t-1)	0.184 (0.551)	0.218 (0.732)	-0.272 (0.996)	-3.056 (2.337)
Number of observations	1592	1195	730	207
Pseudo R-Squared	0.069	0.077	0.103	0.169

Notes: Conditional logit model with year- and district-fixed effects. The dependent variable is conflict onset/ending in (t) operationalized as indicated in the column heading. Standard errors are in parentheses (clustered at the district level).

Effect of Ethnic Fractionalization

Table F.4: No difference conditional on ethnic diversity

Interaction with	Ethnic Groups (GREG)		Ethnic Groups (NRVA)	
	1 if Mixed (1)	Number (2)	1 if Mixed (3)	Number (4)
Opium Profitability (t-1)	-0.763 (0.370)	-0.977 (0.568)	-0.403 (0.380)	-0.380 (0.530)
Opium Profitability (t-1)*X	0.130 (0.421)	0.114 (0.280)	-0.524 (0.423)	-0.179 (0.223)
Number of observations	5174	5174	4615	4615

Notes: Models include province-times-year- and district-fixed effects. The dependent variable is the log of battle-related deaths in (t). Opium Profitability is defined as the interaction between the normalized drug prices (in logarithms) and the suitability to grow opium. Regressions include interactions of the opium profitability with a variable X as indicated in the column heading. For definitions of the variables X please see Appendix A. Standard errors are in parentheses (clustered at the district level).

Modifications of the treatment variable: Drug prices

Table F.5: Non-normalized drug prices, 2002-2014

	(log) BRD (1)	1 if ≥ 5 (2)	1 if ≥ 10 (3)	1 if ≥ 25 (4)	1 if ≥ 100 (5)
Panel A: Local Opium Price					
Opium Profitability (t-1)	-0.644 (0.200)	-0.166 (0.059)	-0.165 (0.056)	-0.143 (0.052)	-0.079 (0.030)
Adjusted R-Squared	0.649	0.501	0.483	0.454	0.311
Panel B: International Heroin Price (baseline)					
Opium Profitability (t-1)	-2.103 (0.835)	-0.503 (0.256)	-0.550 (0.233)	-0.465 (0.206)	-0.183 (0.108)
Adjusted R-Squared	0.649	0.501	0.483	0.453	0.310
Panel C: Complement Price					
Opium Profitability (t-1)	-4.023 (1.337)	-1.016 (0.399)	-0.982 (0.364)	-0.870 (0.329)	-0.371 (0.176)
Adjusted R-Squared	0.650	0.502	0.484	0.454	0.311

Notes: Models include province-times-year- and district-fixed effects. The dependent variable is conflict in (t) operationalized as indicated in the column heading. Opium Profitability is defined as the interaction between the drug prices (in logarithms) and the suitability to grow opium. The number of observations is 5,174 across columns and panels. Standard errors are in parentheses (clustered at the district level).

Table F.6: International heroin price, price not in logarithms, 2002-2014

	(log) BRD (1)	1 if ≥ 5 (2)	1 if ≥ 10 (3)	1 if ≥ 25 (4)	1 if ≥ 100 (5)
Opium Profitability (t-1)	-6.970 (2.232)	-1.665 (0.696)	-1.781 (0.618)	-1.438 (0.537)	-0.553 (0.277)
Adjusted R-Squared	0.649	0.501	0.483	0.453	0.310

Notes: Models include province-times-year- and district-fixed effects. The dependent variable is conflict in (t) operationalized as indicated in the column heading. Opium Profitability is defined as the interaction between the normalized international price (prices are not in logarithms) and the suitability to grow opium. The number of observations is 5,174 across columns. Standard errors are in parentheses (clustered at the district level).

Table F.7: International heroin price, de-trended prices, 2002-2014

	(log) BRD (1)	1 if ≥ 5 (2)	1 if ≥ 10 (3)	1 if ≥ 25 (4)	1 if ≥ 100 (5)
Opium Profitability (t-1)	-6.799 (2.232)	-1.566 (0.681)	-1.882 (0.606)	-1.265 (0.541)	-0.482 (0.248)
Adjusted R-Squared	0.647	0.499	0.481	0.452	0.310

Notes: Models include province-times-year- and district-fixed effects. The dependent variable is conflict in (t) operationalized as indicated in the column heading. Opium profitability is a product of the de-trended heroin price with opium suitability. The number of observations is 5,174 across columns. Standard errors are in parentheses (clustered at the district level).

Table F.8: International heroin price, in deviations, 2002-2014

	(log) BRD (1)	1 if ≥ 5 (2)	1 if ≥ 10 (3)	1 if ≥ 25 (4)	1 if ≥ 100 (5)
Opium Profitability (t-1)	-6.197 (3.136)	-1.434 (0.875)	-1.567 (0.887)	-1.387 (0.747)	-0.620 (0.350)
Adjusted R-Squared	0.649	0.501	0.483	0.453	0.310

Notes: Models include province-times-year- and district-fixed effects. The dependent variable is conflict in (t) operationalized as indicated in the column heading. Opium Profitability is defined as the interaction between international price deviations (from the mean) and the suitability to grow opium. Standard errors are in parentheses (clustered at the district level).

Table F.9: Main results using cocaine prices, 2002-2014 period

	(log) BRD (1)	1 if ≥ 5 (2)	1 if ≥ 10 (3)	1 if ≥ 25 (4)	1 if ≥ 100 (5)
Opium Profitability (t-1)	-0.461 (0.199)	-0.116 (0.059)	-0.124 (0.057)	-0.102 (0.051)	-0.026 (0.025)
Adjusted R-Squared	0.650	0.501	0.484	0.454	0.310

Notes: Models include province-times-year- and district-fixed effects. The dependent variable is conflict in (t) operationalized as indicated in the column heading. Opium Profitability is defined as the interaction between the normalized drug prices (in logarithms) and the suitability to grow opium. The number of observations is equal across all panels. The number of observations is 5,174 across columns. Standard errors are in parentheses (clustered at the district level). Compare with main results in Table 1.

Modifications of the treatment variable: Suitability

Table F.10: Unweighted suitabilities, 2002-2014

	(log) BRD (1)	1 if ≥ 5 (2)	1 if ≥ 10 (3)	1 if ≥ 25 (4)	1 if ≥ 100 (5)
Opium Profitability (t-1)	-0.988 (0.290)	-0.249 (0.093)	-0.244 (0.088)	-0.188 (0.077)	-0.031 (0.041)
Adjusted R-Squared	0.650	0.501	0.483	0.453	0.310

Notes: Models include province-times-year- and district-fixed effects. The dependent variable is conflict in (t) operationalized as indicated in the column heading. Opium Profitability is defined as the interaction between the normalized international prices (in logarithms) and the unweighted suitability to grow opium (in analogy for wheat). The number of observations is 5,174 across columns. Standard errors are in parentheses (clustered at the district level).

Modifications of the treatment variable: Dyadic Diff-in-Diff

Table F.11: Diff-in-Diff - Dyadic treatment, 2002-2014

	(log) BRD (1)	1 if ≥ 5 (2)	1 if ≥ 10 (3)	1 if ≥ 25 (4)	1 if ≥ 100 (5)
Panel A: Suitability dichotomized					
Opium Profitability (t-1)	-0.229 (0.087)	-0.052 (0.027)	-0.041 (0.026)	-0.042 (0.022)	-0.017 (0.013)
Adjusted R-Squared	0.648	0.500	0.482	0.452	0.311
Panel B: Suitability and Heroin Price dichotomized					
Opium Profitability (t-1)	-0.397 (0.117)	-0.107 (0.038)	-0.090 (0.035)	-0.071 (0.030)	-0.029 (0.018)
Adjusted R-Squared	0.649	0.501	0.483	0.453	0.311

Notes: Models include province-times-year- and district-fixed effects. The dependent variable is conflict in (t) operationalized as indicated in the column heading. Suitability (for opium and wheat) dichotomized according to the sample median in panel A. Suitability (for opium and wheat) and international prices (for heroin and wheat) dichotomized according to the sample median in panel B. The number of observations is 5,174 across columns and panels. Standard errors are in parentheses (clustered at the district level).

Adaptions of empirical model

Table F.12: Main results using normalized drug prices, district- and year-fixed effects, 2002-2014

	(log) BRD (1)	1 if ≥ 5 (2)	1 if ≥ 10 (3)	1 if ≥ 25 (4)	1 if ≥ 100 (5)
Panel A: Local Opium Price					
Opium Profitability (t-1)	-0.280 (0.091)	-0.097 (0.028)	-0.078 (0.026)	-0.026 (0.021)	0.000 (0.012)
Adjusted R-Squared	0.562	0.422	0.410	0.389	0.264
Panel B: International Heroin Price (baseline)					
Opium Profitability (t-1)	-0.451 (0.209)	-0.132 (0.064)	-0.122 (0.063)	-0.050 (0.051)	-0.010 (0.023)
Adjusted R-Squared	0.561	0.421	0.409	0.389	0.264
Panel C: Complement Price					
Opium Profitability (t-1)	-0.707 (0.222)	-0.217 (0.068)	-0.172 (0.064)	-0.091 (0.050)	-0.028 (0.023)
Adjusted R-Squared	0.563	0.423	0.410	0.390	0.264

Notes: Models include year- and district-fixed effects. The dependent variable is conflict in (t) operationalized as indicated in the column heading. Opium Profitability is defined as the interaction between the normalized drug prices (in logarithms) and the suitability to grow opium. The number of observations is 5,174 across columns and panels. Standard errors are in parentheses (clustered at the district level).

Table F.13: Population-density-specific heterogeneity, 2002-2014

Interaction with	Continuous (1)	<10 percentile (2)	<20 percentile (3)
Opium Profitability (t-1)	-0.978 (0.295)	-1.019 (0.283)	-1.032 (0.308)
Opium Profitability (t-1)*X	-0.021 (0.075)	0.413 (1.319)	0.198 (0.940)
Adjusted R-Squared	0.649	0.649	0.649

Notes: Models include province-times-year- and district-fixed effects. The dependent variable is the log of battle-related deaths in (t). Regressions include interactions of the opium profitability with different indicators of population density (X) as indicated in the column heading. Column 1 includes a continuous measure on population density, columns 2 and 3 include binary indicators taking a value of one if the population density is below the 10th or 20th percentile. Opium Profitability is defined as the interaction between the normalized international prices (in logarithms) and the unweighted suitability to grow opium (in analogy for wheat). The number of observations is 5,174 across columns. Standard errors are in parentheses (clustered at the district level).

Outcome and timing (cultivation and revenue)

Table F.14: Effect of income shocks on opium cultivation, 2002-2014

	Outcome: (t) (1)	Outcome: (t)+(t-1) (2)
Opium Profitability (t-1)	0.483 (0.307)	0.705 (0.308)
Adjusted R-Squared	0.399	0.488

Notes: The dependent variables opium cultivation is in logarithms. Column (1) presents lagged effects. Column (2) reports lagged and contemporaneous effects by defining the outcome as the moving average, i.e. $(\text{revenues}(t) + \text{revenues}(t-1))/2$. Opium Profitability is defined as the interaction between the normalized heroin price (in logarithms) and the suitability to grow opium. The estimations do not include province-times-year-fixed effects, as the actual district-level cultivation data from which revenues are calculated is gathered at the province level. The number of observations is 5,174 across columns. Standard errors clustered at the district level are displayed in parentheses.

Standard errors

Table F.15: Standard errors clustered at different levels, 2002-2014

	(log) BRD (1)	1 if ≥ 5 (2)	1 if ≥ 10 (3)	1 if ≥ 25 (4)	1 if ≥ 100 (5)
Panel A: Clustered at district and year level					
Opium Profitability (t-1)	-0.675 (0.325)	-0.167 (0.092)	-0.191 (0.080)	-0.147 (0.074)	-0.040 (0.051)
Adjusted R-Squared	0.649	0.501	0.484	0.453	0.310
Panel B: Clustered at province and year level					
Opium Profitability (t-1)	-0.675 (0.365)	-0.167 (0.103)	-0.191 (0.103)	-0.147 (0.082)	-0.040 (0.044)
Adjusted R-Squared	0.649	0.501	0.484	0.453	0.310

Notes: Models include province-times-year- and district-fixed effects. The dependent variable is conflict in (t) operationalized as indicated in the column heading. The number of observations is 5,174 across columns and panels. Standard errors are clustered as indicated in the panel heading.

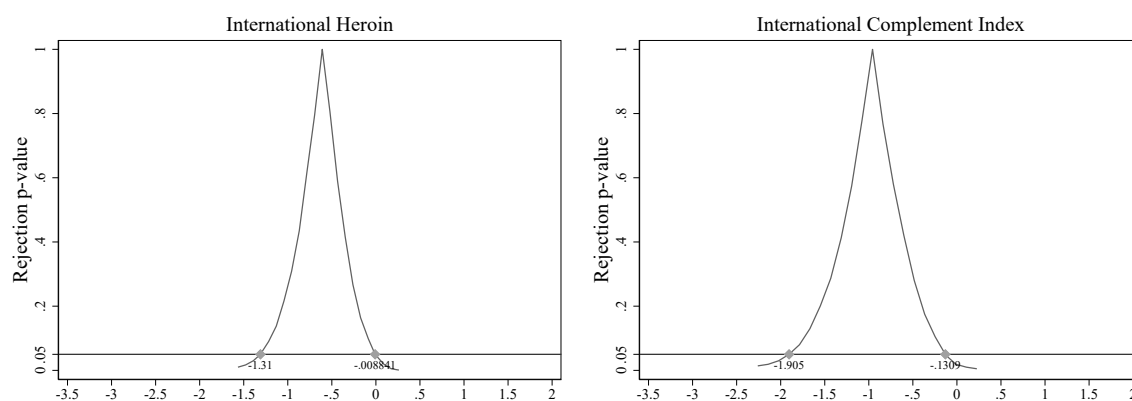


Figure F.3: Wild Bootstrap (province level clustered se, 95% confidence intervals)

Notes: Figures show the distribution of bootstrap estimates. The dependent variable is the (log) of BRD. Regressions correspond to Table 1 column 1 (panel B and C). The number indicate the left and right 95%-confidence interval. The test of the the null hypothesis at the 5%-level is whether this intervall contains 0.

Covariates and trends

Table F.16: Wheat Profitability, 2002-2014

	(log) BRD (1)	1 if ≥ 5 (2)	1 if ≥ 10 (3)	1 if ≥ 25 (4)	1 if ≥ 100 (5)
Panel A: Wheat Profitability included					
Opium Profitability (t-1)	-0.675 (0.296)	-0.167 (0.090)	-0.191 (0.085)	-0.147 (0.075)	-0.040 (0.037)
Wheat Profitability (t-1)	0.307 (0.123)	0.088 (0.039)	0.077 (0.036)	0.034 (0.031)	-0.010 (0.019)
Adjusted R-Squared	0.649	0.501	0.484	0.453	0.310
Panel B: Wheat Profitability excluded					
Opium Profitability (t-1)	-0.923 (0.279)	-0.238 (0.085)	-0.253 (0.079)	-0.175 (0.069)	-0.031 (0.030)
Adjusted R-Squared	0.649	0.501	0.483	0.453	0.310

Notes: Models include province-times-year- and district-fixed effects. The dependent variable is conflict in (t) operationalized as indicated in the column heading. The number of observations is 5,174 across columns and panels. Standard errors are in parentheses (clustered at the district level).

Table F.17: Dynamics - lagged dependent, 2002-2014

	(log) BRD (1)	1 if ≥ 5 (2)	1 if ≥ 10 (3)	1 if ≥ 25 (4)	1 if ≥ 100 (5)
Opium Profitability (t-1)	-0.455 (0.252)	-0.140 (0.084)	-0.160 (0.076)	-0.102 (0.065)	-0.021 (0.032)
Dependent (t-1)	0.236 (0.023)	0.114 (0.019)	0.153 (0.023)	0.228 (0.027)	0.207 (0.040)
Adjusted R-Squared	0.670	0.508	0.495	0.481	0.340

Notes: Models include province-times-year- and district-fixed effects. The dependent variable is conflict in (t) operationalized as indicated in the column heading. The number of observations is 5,174 across columns. Standard errors are in parentheses (clustered at the district level).

Table F.18: Including covariates, 2002-2014

	(log) BRD (1)	1 if ≥ 5 (2)	1 if ≥ 10 (3)	1 if ≥ 25 (4)	1 if ≥ 100 (5)
Panel A: Baseline covariates					
Opium Profitability (t-1)	-0.595 (0.275)	-0.177 (0.086)	-0.188 (0.082)	-0.132 (0.070)	-0.014 (0.038)
Wheat Profitability (t-1)	0.282 (0.129)	0.093 (0.041)	0.077 (0.037)	0.028 (0.032)	-0.019 (0.019)
VHI (t)	0.000 (0.002)	-0.000 (0.001)	-0.000 (0.001)	0.000 (0.000)	-0.000 (0.000)
Luminosity (t-2)	0.018 (0.020)	0.005 (0.006)	0.002 (0.006)	-0.003 (0.005)	-0.000 (0.003)
(log) Population (t-2)	1.417 (3.472)	-0.478 (0.911)	0.037 (0.900)	0.611 (0.959)	0.789 (0.308)
Number of observations	5173	5173	5173	5173	5173
Adjusted R-Squared	0.649	0.501	0.483	0.453	0.311
Panel B: Baseline covariates, time-invariant covariates×trend					
Opium Profitability (t-1)	-0.645 (0.270)	-0.166 (0.084)	-0.195 (0.081)	-0.179 (0.068)	-0.031 (0.040)
Wheat Profitability (t-1)	0.285 (0.132)	0.091 (0.042)	0.079 (0.038)	0.024 (0.032)	-0.015 (0.020)
VHI (t)	-0.000 (0.002)	-0.000 (0.001)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)
Luminosity (t-2)	0.014 (0.019)	0.004 (0.006)	0.001 (0.006)	-0.004 (0.005)	-0.000 (0.003)
(log) Population (t-2)	-0.300 (3.443)	-0.712 (0.954)	-0.646 (0.948)	-0.219 (0.916)	0.877 (0.387)
Number of observations	5147	5147	5147	5147	5147
Adjusted R-Squared	0.654	0.504	0.488	0.462	0.317
Panel C: Baseline covariates, time-invariant covariates×time dummies					
Opium Profitability (t-1)	-0.757 (0.288)	-0.207 (0.088)	-0.223 (0.086)	-0.195 (0.073)	-0.045 (0.043)
Wheat Profitability (t-1)	0.247 (0.143)	0.079 (0.042)	0.072 (0.041)	0.022 (0.036)	-0.024 (0.022)
VHI (t)	-0.000 (0.002)	-0.001 (0.001)	-0.000 (0.001)	0.000 (0.000)	-0.000 (0.000)
Luminosity (t-2)	0.010 (0.020)	0.002 (0.006)	-0.000 (0.006)	-0.005 (0.005)	-0.000 (0.003)
(log) Population (t-2)	-0.762 (3.604)	-0.795 (1.001)	-0.793 (0.988)	-0.375 (0.956)	0.843 (0.391)
Number of observations	5147	5147	5147	5147	5147
Adjusted R-Squared	0.654	0.503	0.487	0.462	0.315

Notes: Models include province-times-year- and district-fixed effects. The dependent variable is conflict in (t) operationalized as indicated in the column heading. The set of time-invariant covariates includes Ruggedness (categorized in five quantiles), Distance Linear, Distance 2D and 3D, Travel Time 2D and 3D (all distances to Kabul), Pashtuns, Mixed Ethnic Groups, Number of Ethnic Groups, Taliban Territory 1996, Mixed Territory 1996, and Ethnic Trafficking Route. Standard errors are in parentheses (clustered at the district level).

Table F.19: Production and opium-suitability-specific trends, 2002-2014

Percentile	Based on production			Based on opium suitability		
	50 (1)	75 (2)	90 (3)	50 (4)	75 (5)	90 (6)
Opium Profitability (t-1)	-0.505 (0.207)	-0.528 (0.210)	-0.489 (0.209)	-0.313 (0.367)	-0.675 (0.296)	-0.675 (0.296)
Adjusted R-Squared	0.564	0.562	0.562	0.650	0.649	0.649

Notes: Models include province-times-year- and district-fixed effects as well as production- or suitability-specific time trends based on the percentile as indicated in the column heading. The dependent variable is the log of battle-related deaths in (t). The number of observations is 5,174 across columns. Standard errors are in parentheses (clustered at the district level).

Outlier analysis

Table F.20: Drop potential outliers, 2002-2014

	(log) BRD (1)	1 if ≥ 5 (2)	1 if ≥ 10 (3)	1 if ≥ 25 (4)	1 if ≥ 100 (5)
Panel A: No border districts					
Opium Profitability (t-1)	-0.601 (0.304)	-0.160 (0.098)	-0.161 (0.096)	-0.146 (0.086)	-0.014 (0.055)
Number of observations	3718	3718	3718	3718	3718
Adjusted R-Squared	0.678	0.523	0.512	0.483	0.341
Panel B: No Southern provinces (Kandahar and Hilmand)					
Opium Profitability (t-1)	-0.674 (0.311)	-0.174 (0.096)	-0.215 (0.091)	-0.118 (0.078)	-0.007 (0.033)
Number of observations	4732	4732	4732	4732	4732
Adjusted R-Squared	0.620	0.479	0.458	0.407	0.254

Notes: Models include province-times-year- and district-fixed effects. The dependent variable is conflict in (t) operationalized as indicated in the column heading. In panel A, all border districts are excluded and in panel B all districts in the two provinces Kandahar and Hilmand are excluded. Standard errors are in parentheses (clustered at the district level).

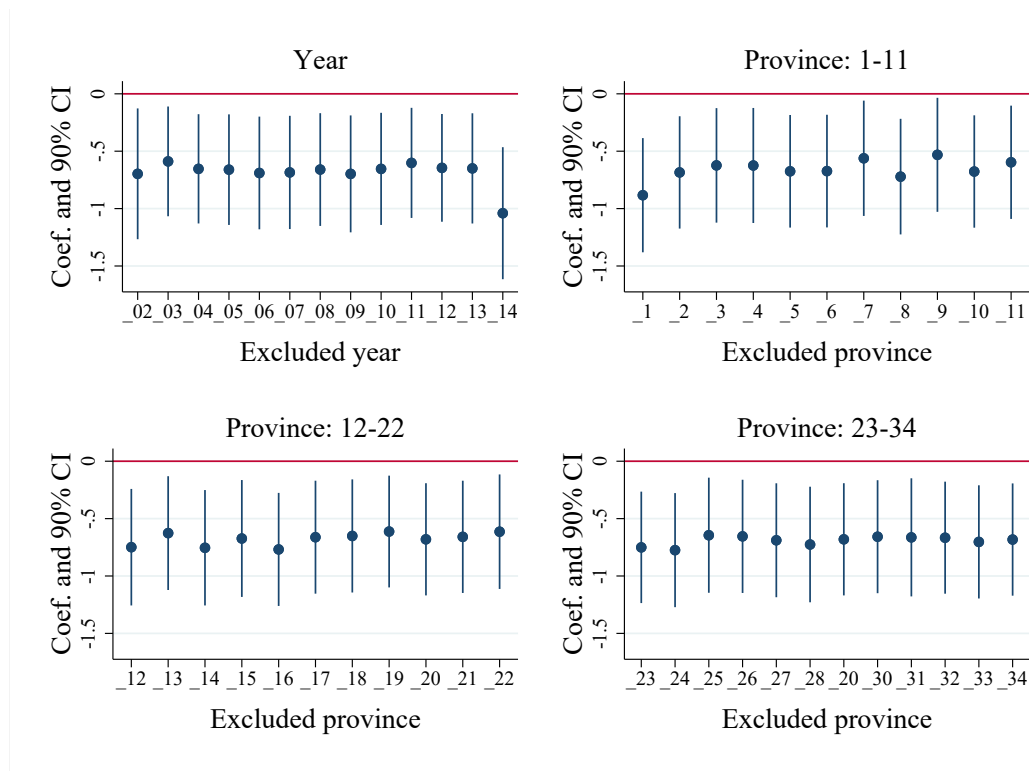


Figure F.4: Leave one out - year and province

Notes: This figure shows results for 47 separate regressions in analogy to panel B's column (1) of Table 1, where we leave out one year or one province at the time. This also alleviates concerns whether particular outliers in the cross-sectional variation drive our result.

Randomization

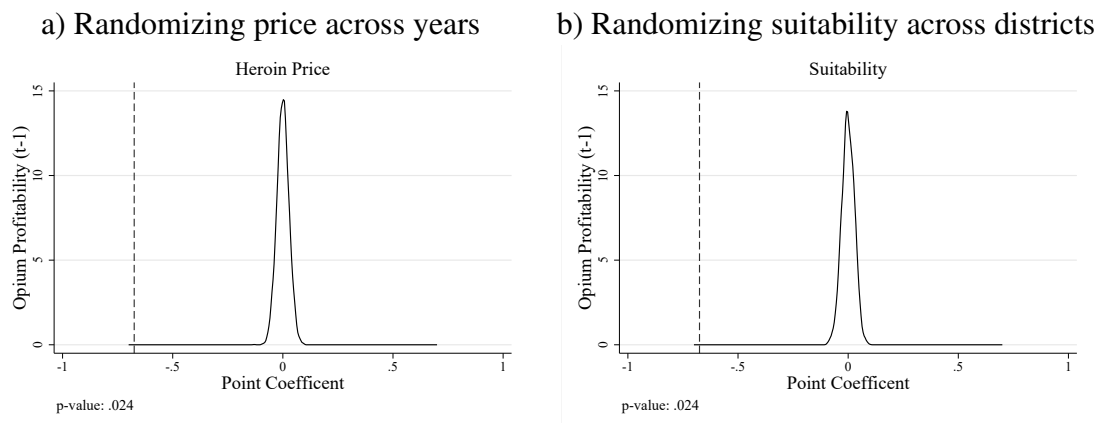


Figure F.5: Randomization: Heroin Price and Opium Suitability

Notes: This figure plots the distribution of the coefficients generated by 5'000 randomizations, with panel A randomly reordering prices across years within districts and multiplying with the actual suitability and panel B reordering the suitability across districts and multiplying with the actual price in the respective yes. Based on the regression model in panel B's column (1) of Table 1. For this placebo test, we want to see whether the randomized coefficients are centered around zero, and what share of the draws turn out to be more negative than the actual treatment coefficient. This share is used to compute the randomization inference p-value shown in the bottom of the graph.

Tables and robustness for regressions at the household level

Table F.21: Living standards at the household level, 2005-2012

	(1)	(2)	(3)
Panel A: Food consumption			
	Dietary Diversity	Calorie Intake	Food Insecurity
Opium Profitability (t-1)	0.410 (0.207)	0.105 (0.187)	0.527 (0.229)
Number of observations	72224	71634	72643
Adjusted R-Squared	0.371	0.139	0.225
Panel B: Food expenditures			
	Food Exp.	Food Exp. Paasche adj.	Food Exp. Laspeyres adj.
Opium Profitability (t-1)	0.527 (0.229)	0.621 (0.246)	0.590 (0.247)
Number of observations	72643	72643	72635
Adjusted R-Squared	0.225	0.196	0.217
Panel C: Assets			
	Sum of Assets	Sum of Assets weighted	Economically improved
Opium Profitability (t-1)	0.580 (0.205)	0.556 (0.197)	0.435 (0.226)
Number of observations	72447	66620	70670
Adjusted R-Squared	0.323	0.336	0.249

Notes: Models include province-times-year- and district-fixed effects. The dependent variable in (t) is operationalized as indicated in the column heading. Outcomes are standardized to have a mean value of zero and a standard deviation of one. Opium Profitability is defined as the interaction between the normalized drug prices (in logarithms) and the suitability to grow opium. Standard errors are in parentheses (clustered at the district-year level).

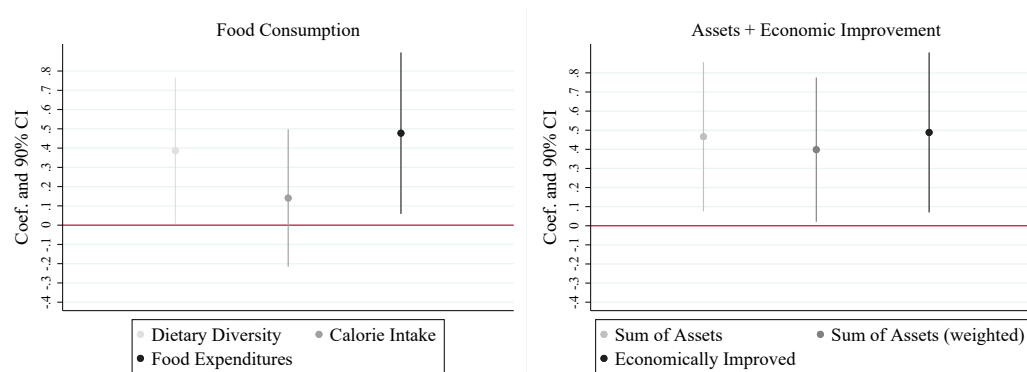


Figure F.6: Effect of Opium Profitability (t-1) on living standard indicators in (t), accounting for household survey weights

Notes: The figure shows results of six separate regressions in analogy to Table F.21. The difference is that we include household survey weights in the regressions. Outcomes are standardized to have a mean value of zero and a standard deviation of one. Results are also robust to using robust standard errors rather than clustering at the district-year level.

Survey results conditional on degree of group competition - focusing on Kabul area The area within 50 km around Kabul is probably the one where the Afghan government has the best means and highest likelihood to enforce laws. Enforcement means more eradication and seizing of opium production. Hence, we would predict that the effect of a higher opium profitability on household food consumption and assets should be less positive. The results in Figure F.7 show that this is the case.

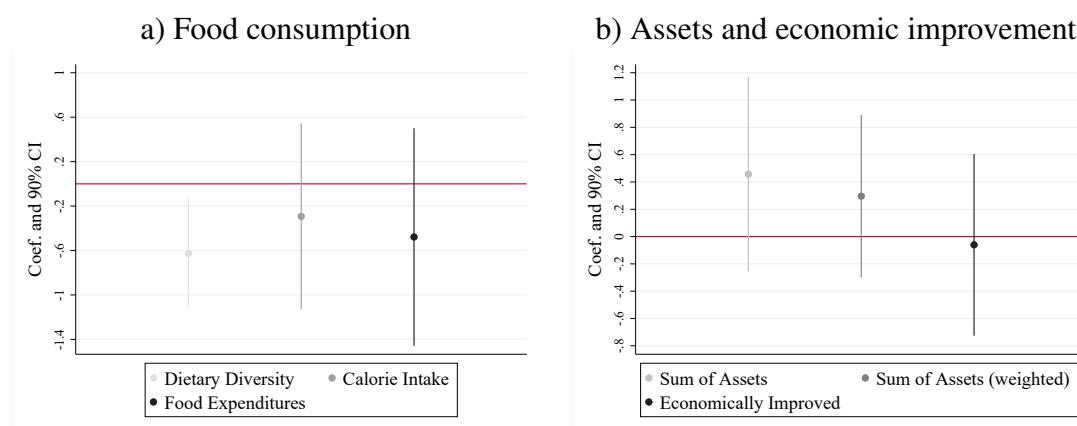


Figure F.7: Effect of Opium Profitability (t-1) on standard of living indicators in (t) close to Kabul

Notes: The figure plots coefficient estimates (along with 90% confidence intervals) of opium profitability for six separate regressions with outcomes as indicated in the legends. The outcome variables are based on the NRVA household survey, details on their construction can be found in Section A. Outcomes are standardized to have a mean value of zero and a standard deviation of one. All regressions follow the baseline regression equation with international heroin prices in Table 1, panel B. The regressions include districts within a 50km distance of Kabul

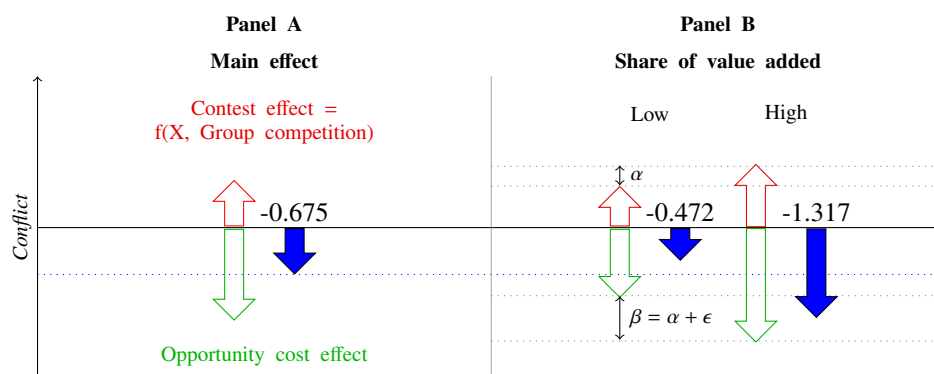


Figure F.8: Opportunity cost effects dominate contest effects (2002-2014)

Notes: Panel A: Opportunity cost effects dominate contest effects on average. Refers to Table 1, panel B, column 1. Panel B: Opportunity costs effects increase more in the share of value added than contest effects. Refers to Table 2, panel A, column 1.

Graphical illustration for heterogeneity across value added, Table 2

Robustness for Table 3

Table F.22: Government control and law enforcement, 2002-2014 period

Proximity to	Military bases		Kabul		Other cities	
Interaction with	Linear distance					
	1 if ≤ 100	1 if ≤ 125	1 if ≤ 100	1 if ≤ 125	1 if ≤ 100	1 if ≤ 125
	(1)	(2)	(3)	(4)	(5)	(6)
Opium Profitability (t-1)	-1.115 (0.431)	-0.963 (0.478)	-0.782 (0.313)	-0.650 (0.329)	-0.530 (0.344)	-0.545 (0.361)
Opium Profitability (t-1)*X	0.896 (0.489)	0.440 (0.527)	0.712 (0.667)	0.334 (0.621)	-0.485 (0.457)	-0.366 (0.407)
Adjusted R-Squared	0.650	0.650	0.650	0.651	0.649	0.649

Notes: Models include province-times-year- and district-fixed effects. The dependent variable is the log of battle-related deaths in (t). Opium Profitability is defined as the interaction between the normalized international heroin price (in logarithms) and the suitability to grow opium. Regressions include interactions of the opium profitability with a variable X as indicated in the column heading. Other cities are Kandahar, Kunduz, Jalalabad, Hirat, and Mazari Sharif (next five largest cities). The number of observations is 5,174 across columns. Standard errors are in parentheses (clustered at the district level).

Table F.23: Alternative proxy for Taliban control (based on NRVA)

Interaction with	Any Pashtuns (1)	Share Pashtuns (2)
Opium Profitability (t-1)	-0.062 (0.403)	-0.283 (0.359)
Opium Profitability (t-1)*X	-1.157 (0.433)	-1.005 (0.577)
Adjusted R-Squared	0.646	0.647

Notes: Models include province-times-year- and district-fixed effects. The dependent variable is the log of battle-related deaths in (t). Opium Profitability is defined as the interaction between the normalized drug prices (in logarithms) and the suitability to grow opium. Regressions include interactions of the opium profitability with a variable X as indicated in the column heading. For definitions of the variables X please see Appendix A. For this table the different measures on ethnic groups are derived from the NRVA 2003, which is not nationally representative, but serves as suitable a proxy for ethnic group distribution. The number of observations is 5,174 across columns. Standard errors are in parentheses (clustered at the district level).

Table for Figure 5

Table F.24: Regressions results corresponding to Figure 5

Measure of Taliban control:	Former Taliban area (1)	Pashtun (Ethnologue) (2)	Pashtun (GREG) (3)
Contested	-0.458 (0.441)	-0.462 (0.412)	0.006 (0.427)
Government Control	0.127 (0.381)	0.173 (0.377)	0.129 (0.371)
Taliban Control	-1.592 (0.449)	-2.195 (0.477)	-1.745 (0.358)

Notes: The dependent variable is the log of battle-related deaths in (t). Standard errors are in parentheses (clustered at the district level). Which district is associated with which scenario is based on the conditions defined in the paper about Taliban or government control. The coefficients are the marginal effects of Opium Profitability conditional on being in one of those three scenarios. Opium Profitability is the interaction between the normalized drug prices (in logarithms) and the suitability to grow opium. The number of observations is 5,174 across columns. All regressions include province-times-year- and district-fixed effects.

G. Maps

a) Opium suitability (based on Kienberger *et al.* , 2017)

b) Wheat suitability (based on FAO GAEZ)

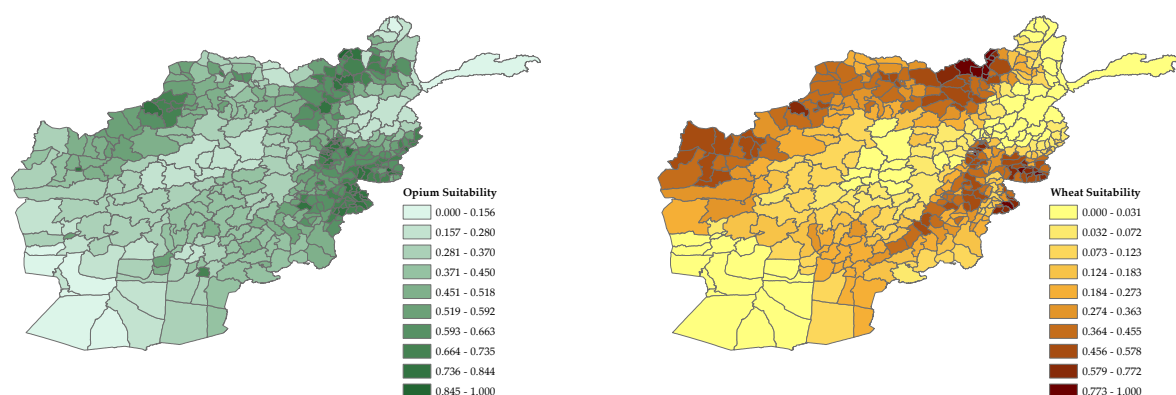


Figure G.1: Distribution of opium and wheat suitability across districts (weighted by population)

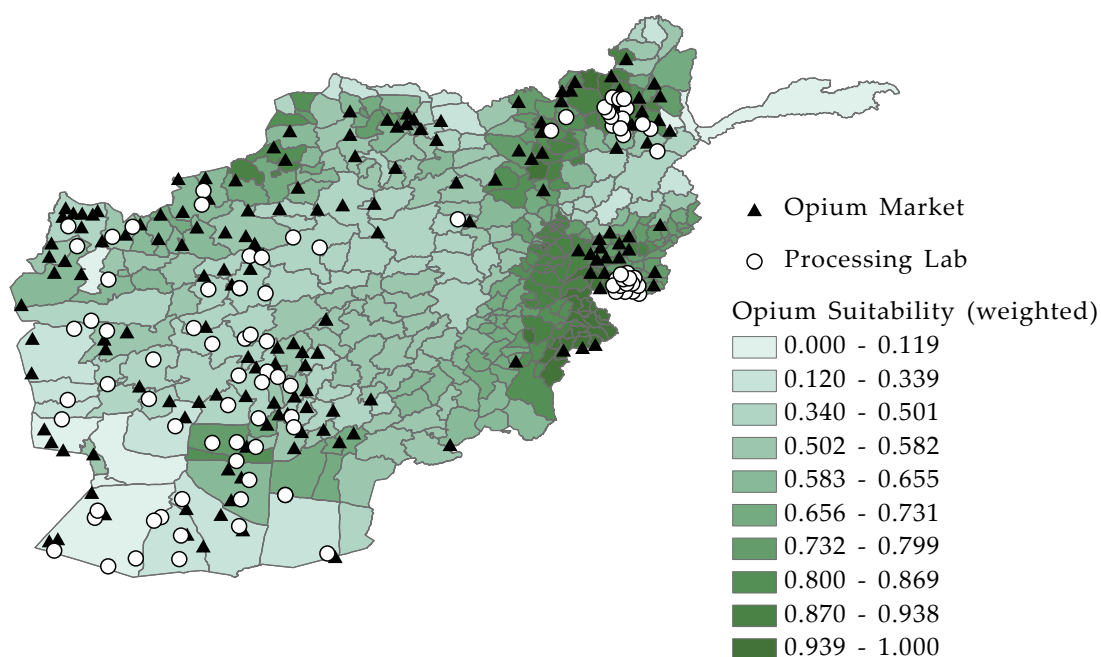


Figure G.2: Opium suitability and the location of opium markets and processing labs

Notes: Opium suitability is from Kienberger *et al.* (2017) and is weighted by population. Opium market and lab information based on UNODC.

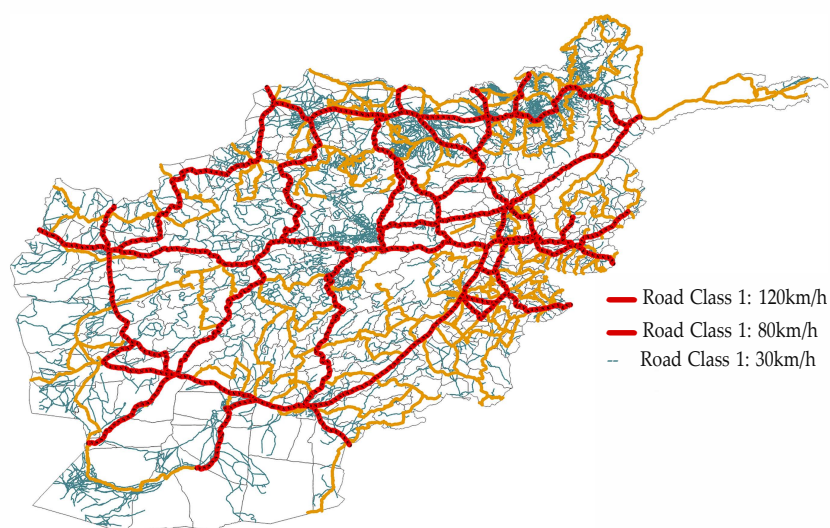


Figure G.3: The road network

Notes: The road network in Afghanistan distinguishing in highways (assumed speed 120 km/h), rural roads (ass. speed 90 km/h), and urban roads (ass. speed 50 km/h). The distinction in road types and the choice of average speed is not decisive for our results.

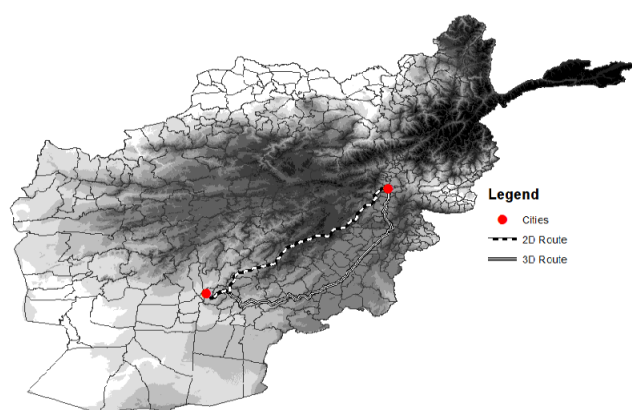


Figure G.4: Elevation and distance

Notes: The intensity of black indicates the elevation in Afghanistan. The white-black dashed line shows the shortest road distance between to district centroids. The second white/black line indicates the shortest distance when accounting for elevation differences along the roads. In particular the central part of Afghanistan is very mountainous, which can have a large effect on transportation costs and travel time.

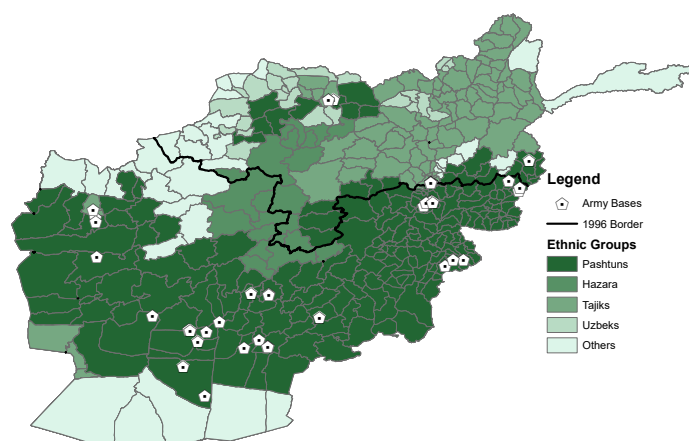


Figure G.5: Distribution of main ethnic groups

Notes: The map shows the four major ethnic groups in Afghanistan in different shades of green (Source: GREG). The white symbols with the black dots indicate the location of a foreign military base, for which we could track location, opening and closing date (sources in detail in Appendix H). The area south of the thick black line was controlled by the Taliban prior to 2001 (Dorronsoro, 2005).

Taliban control: 1996 border

The figure below is an excerpt from a book by Dorronsoro (2005). We geo-reference the green area as the area formerly under Taliban control, and the other three polygons as not under Taliban control.

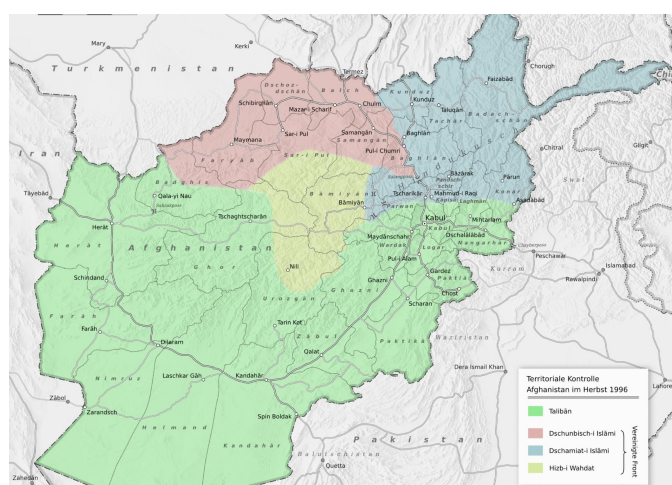


Figure G.6: Political control in Afghanistan in the fall of 1996

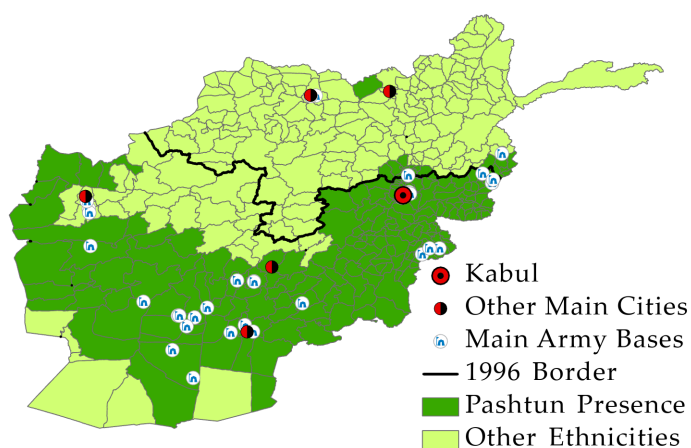


Figure G.7: Pashtun presence (Ethnologue), main army bases, and major cities

Notes: Presence of the ethnic group Pashtun (Source: Ethnologue). Red dots indicate the capital Kabul and other major cities: Hirat, Kandahar, Kunduz, Mazari Sharif. White symbols with black dots indicate the location of a foreign military base, for which we could track location, opening and closing date (sources in detail in Appendix H). The area south of the thick black line was controlled by the Taliban prior to 2001 (Dorrnsoro, 2005).

Taliban control: Pashtun presence based on GREG dataset

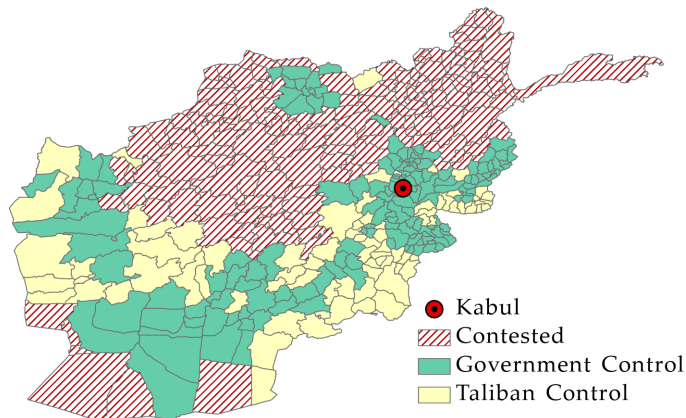
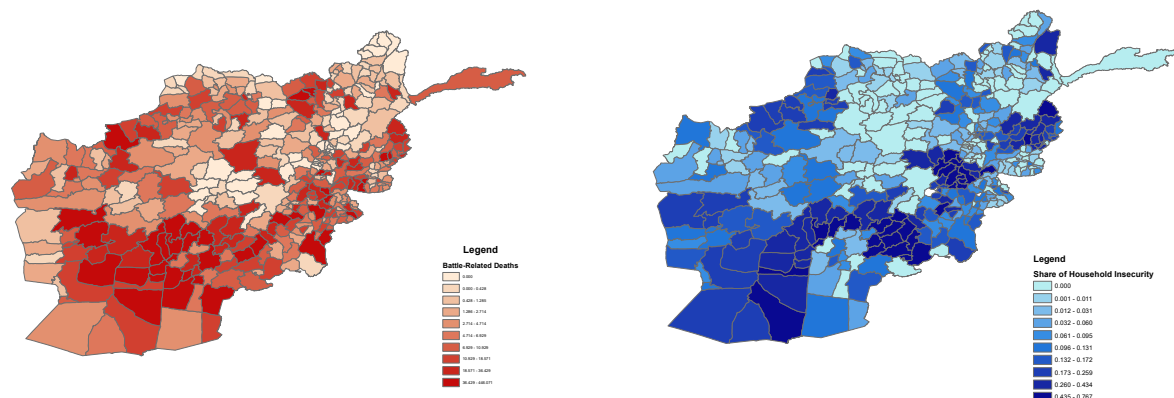


Figure G.8: Control across districts (Pashtun, based on Ethnologue)

Notes: See Figure 5 for the same map using former Taliban territory to determine Taliban control.



Number of battle-related deaths (UCDP GED)

Share of households experiencing insecurity shock (NRVA)

Figure G.9: Distribution of objective (BRD) and subjective (NRVA) conflict indicators (2002-2014)

H. Data coding and map generation

Processing and trafficking: There is little to no information that is publicly available on trafficking routes that might be used to smuggle opium through and out of the country. Nevertheless, the UN Office on Drugs and Crime creates and contains spatial maps in its public reports. We digitize a UNODC map from 2007 (about the middle of our sample period) by taking image files of the maps themselves and georeferencing specific points on the images (border points) to a geographically accurate projection of Afghanistan. This process was continued until the map and the images matched perfectly. We then digitized the data contained in the image about the important roads used for trafficking, and the other variables such as main opium markets and heroin processing labs. We verified locations with other UNODC reports.

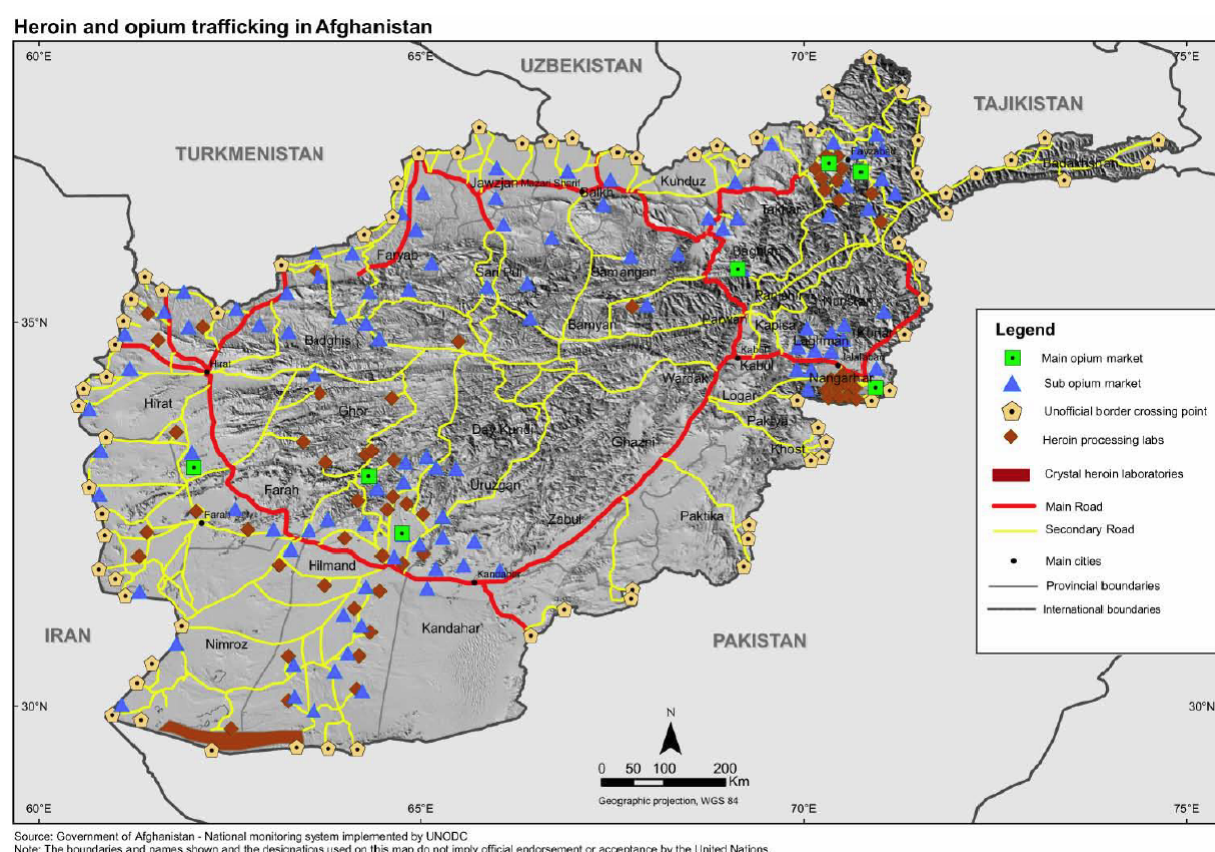


Figure H.1: Original UNODC map (2007)

Map making process: The source of the original map is UNODC's 2007 Afghanistan Opium Survey. The map depicts major and secondary roads, main cities, opium markets, border crossing points, and processing labs. We also used the 2009 Afghanistan Opium Survey

to cross-validate the data points. In almost all cases, there were no changes between the two years. In case the 2009 map identifies additional markets or labs we added these as data points. Given that the location of illegal markets and labs will always contain some measurement error and could be moved over time, our aim is to code variables that measure the potential for a trafficking route, border crossing, market or lab. This means that the indicators that we create are time-invariant, also due to the availability of data. We interact the binary indicators extracted from the map with an exogenous variable, so that the interaction term can be interpreted as causal under relatively mild circumstances.



Figure H.2: Superimposed maps

In the next step of the process, we match the borders of the image and the geo-referenced (Coordinate system GCS WGS 1984) shapefile for Afghan authorities.⁸ This way, we are accurately overlaying the data points and not simply making an educated guess as to where to place the points. Below are the two final digitized maps based on the UNODC data, overlaid with the district data. The binary indicators that we use in Section 5 on heterogeneous effects are coded as one if the respective feature is present within the boundaries of the district polygon at least once. Alternatively, we use the number of feature per district, e.g., for opium markets.

Major military bases: This section describes how we determine the locations of major known military bases in Afghanistan. There are nearly 400 foreign military bases in Afghanistan, but most bases release no official information as to their geographic location for security rea-

⁸ ESOC Princeton, <https://esoc.princeton.edu/data/administrative-boundaries-398-districts>, last accessed December 09, 2021.

sons. In order to find this information, we compile data from different sources about the most relevant bases to include. We then pinpoint, with latitude and longitude coordinates, the exact location of these bases. Since some are now closed, some data points record past base locations. We rely on information from Wikipedia's GeoHack program for the more well-known bases and on news articles, Wikimapia and Google Maps satellite data for the less well-documented ones. News articles were useful in this case because they are often allowed to publish the district in which these bases are located; from there, we were able to look for these bases by referencing photos of the bases (if available) with available satellite data to verify their location. Below, we show the table with the locations of the about 50 bases that we could identify. The exact locations are blackened out for confidentiality reasons, even though we are convinced none of this information is confidential and could be misused or endanger soldiers. Without access to confidential NATO and US military information this is the best data we could assemble. It is certainly not a complete list of bases, which introduces considerable measurement error to the indicator variable we create based on it. At the same time, we have no reason to expect this measurement error to be non-normal.

OBJECTID *	Base Name	Installation Type	Militaries Present	Lat	Lon	District
1	Delaram	FOB	USMC			Delaram
2	Lasharack	Camp	USMC			Nahri Saraj
3	Kabul International Airport	Camp	ISAF, Turkish Army, US Army, USMC, USAF, Mongolian Armed Forces			Kabul
4	Kandahar Airfield	Airfield	ISAF, USAF, US Army			Kandahar
5	Shindand Airbase	Airbase	USAF, USAF			Shindand
6	Bagram Airfield	Airfield	US Army, USAF			Bagram
7	Basoon	Camp	British Army, RAF, Royal Navy (RN), Royal Marines (RM), USMC, Estonian Land Forces, Danish Defence, Tonga Defence Services			Nahri Saraj
8	Pirca	FOB	Rt British Army, Danish Defence, US Army, USMC			Nahri Saraj
9	Lashar Gah	FOB	British Army, RM			Lashargah
10	Eggers	Camp	NATO, US Army, USMC, US Air Force, Australian Army, New Zealand Army, French Army, Turkish Army, Mongolian Armed Forces			Kabul
11	Sakono	FOB	US Army, USAF, US Navy			Khost (Matun)
12	Chapman	FOB	US Special Operation Command, US Army, CIA			Khost (Matun)
13	Marmal	Camp	German Army, German Navy, German Air Force, Royal Netherlands AF, Swedish Air Force, US Army, Mongolian Armed Forces			Wazir-e Sharif
14	Deyar	Camp	USMC, British Army, RM			Garmar
15	Rhino	Camp	USMC, US Navy, US Army, USAF, SASR			Garmar
16	Holand	Camp	Australian Army, New Zealand Army, US Army, Royal Netherlands Army, ANA			Tarin Kot
17	Black Horse	Camp	US Army, Canadian Army			Kabul
18	Dogan	Camp	<Null>			Kabul
19	Invids	Camp	Italian Army			Kabul
20	Julen	Camp	Canadian Army			Kabul
21	Julen	Camp	Canadian Army			Kabul
22	Phoenix (Qargha)	Camp	US Army			Kabul
23	Soulen	Camp	British Army			Kabul
24	Warehouse	Camp	Canadian Army			Kabul
25	Pucio	Camp	USOCOM			Khost (Matun)
26	Clark	Camp	US Army			Handouza
27	Blessing	Camp	US Army, USMC			Yuliyar
28	Boosick	FOB	US Army			Nari
29	Joyce	FOB	US Army			Sarkani
30	Vingrit	Camp	US Army			Asadabad
31	Albert	Camp	US Army			Bagram
32	Blackjack	Camp	US Army			Bagram
33	Bulldog	Camp	US			Bagram
34	Civilian	Camp	US			Bagram
35	Cunningham	Camp	US			Bagram
36	Gibraltar	Camp	US			Bagram
37	Warrior	Camp	US			Bagram
38	Patt	Camp	US Army			Hazar-e Sharif
39	Spann	Camp	US Army			Hazar-e Sharif
40	Balan	Camp	Australian Army			Daman
41	Hathan Smith	Camp	Canadian Army, US Army			Kandahar
42	Hadian	Camp	Royal Netherlands Army			Del Rawod
43	Russell	Camp	Australian Army			Tarin Kot
44	Hamasulan	FOB	USMC, British Army, RM			Sangin
45	Arena	Camp	Italian Army, Italian Air Force, US Army			Hirat
46	Stone	Camp	Carabinieri, US Army			Hirat
47	Vanni	Camp	Italian Army			Hirat
48	Losano	Camp	RtLAF, US Army, USAF			Kandahar
49	Lagman	FOB	US Army, US Navy, Romanian Army, ANA			Qalat
50	Shorabak	FOB	ISAF, US, Britain, Denmark, Estonia, Tonga			Lashargah
51	Passe (Wilson)	FOB	US Army			Panjwayi

Figure H.3: Main bases and relevant information (1/2)

	Opened	Closed	Field9	Notes	Shape *
2009	2014	<Null>		<Null>	Point
2008	2014	<Null>		Regional Command Southwest Headquarters	Point
2001		Open		ISAF Headquarters, ISAF Joint Command Headquarters, Headquarters for RC-Capital	Point
2001		Open		RC-S Headquarters	Point
2004	2014	<Null>		<Null>	Point
2001		Open		Largest US base in Afghanistan, RC-East Headquarters	Point
2006	2014	<Null>		Main British base and formerly home to Task Force Helmand	Point
2006	2014	<Null>		<Null>	Point
2006	2014	<Null>		NATO Training Mission - Afghanistan Headquarters	Point
2003	2013	<Null>		<Null>	Point
2001		Open		Major CIA and Special Operations counter-insurgency outpost	Point
2005		Open		<Null>	Point
2007	2009	<Null>		<Null>	Point
2001	2002	<Null>		First Marine land base in Afghanistan	Point
2006	2013	<Null>		<Null>	Point
2006	2013	<Null>		<Null>	Point
2002	2015	<Null>		<Null>	Point
2006	2012	Close unk, camp was open in 2012		<Null>	Point
2003	2005	<Null>		Reopened as a Counterinsurgency Academy in April 2007	Point
2007		Open		Reopened as a Counterinsurgency Academy in April 2008	Point
2007		Open		Opening unknown	Point
2007	2014	Stated to close in 2014		<Null>	Point
2002	2014	Stated to close in 2014, Canada withdrew all troops at this time		<Null>	Point
2002	2013	<Null>		<Null>	Point
2002		Open unk, close unk		<Null>	Point
2002	2011	<Null>		<Null>	Point
2006	2012	<Null>		<Null>	Point
2002	2013	Close unk, camp was open in 2013		<Null>	Point
2001		Close unk		<Null>	Point
2004	2012	Close unk, camp still open 2012		Located in related to Bagram Airfield	Point
2004	2012	Open unk, close unk, camp still open 2012		Located in related to Bagram Airfield	Point
2003	2012	Close unk, camp still open 2012		Located in related to Bagram Airfield	Point
2004	2012	Close unk, camp still open 2012		Located in related to Bagram Airfield	Point
2002	2012	Close unk, camp still open 2012		Located in related to Bagram Airfield	Point
2002	2012	Close unk, camp still open 2012		Located in related to Bagram Airfield, opening date unknown	Point
2002	2014	Open unk		<Null>	Point
2002	2014	Open unk, between 2001 and 2004		<Null>	Point
2006	2015	<Null>		Located in related to Kandahar Airfield	Point
2003	2013	<Null>		<Null>	Point
2003	2013	Open unk, task force Urugzan started 2006		<Null>	Point
2005	2013	<Null>		<Null>	Point
2007	2014	<Null>		<Null>	Point
2012		Open		<Null>	Point
Before 2008	2014	<Null>		<Null>	Point
Before 2008	2012	<Null>		Located in related to Bagram Airfield	Point
2004	2014	Open unk, close unk		Located in related to Kandahar Airfield	Point
2005	2014	<Null>		<Null>	Point
2005	Open			ISAF logistics hub	Point
2005	2014	Open unk, slated to close in 2014		<Null>	Point

Figure H.4: Main bases and relevant information (2/2)

This table shows the available data for about 50 bases that we deemed to be the most important foreign bases in Afghanistan over the last 15 years. We list the name, type, location (coordinate system CGS WGS 1984), district, troop's origin, date opened and closed (a “.” in the opened or closed section means there is either no data for closure time or that the base is still open. See Field 9 for explanatory notes in these cases), and general notes of interest.



Figure H.5: Confirming the location of these districts using satellite imagery
Example: “Base Blackhorse”

This is an example of the Wikimapia satellite imagery, which we used to locate bases. This is an image of Base Blackhorse, now closed. We were able to locate this as Base Blackhorse by first searching for the military base on wikimapia which offered two possible locations (approximately 9 miles away from each other) where the base could be. After we discovered in a news report that the base was located next to an Afghan National Army base, which was itself located on the site of the Pul-e-Charkhi-Prison, we were able to determine the definitive location of the prison and thus the location of the base.

Definitions and explanation of how each base was found. Below, we have laid out the definitions for what each type of base exists in Afghanistan and explained how we determined the specific locations for each base we included. The base definitions are important to know because the type of base is a good indicator of its size. Though this was not the only criteria we used to determine whether or not a specific base should be represented on the map, it was important for weeding out those that are not included (for example, we included no firebases on account of their temporary nature and generally small size). Below this, we provide more detail about specific bases whose locations we were not able to get from the GeoHack database,

in which bases are supposed to have had multiple confirmations. These bases were found using satellite data and through available news reports, photos and satellite imagery. All definitions below are adapted or directly taken from Wikipedia to provide a rough idea about the types of military bases that exist in Afghanistan. We simply code whether there is an open base or not.

Additional information about bases (from wikipedia):

- **Definition FOB** - A forward operating base (FOB) is any secured forward military position, commonly a military base, that is used to support tactical operations. A FOB may or may not contain an airfield, hospital, or other facilities. The base may be used for an extended period of time. FOBs are traditionally supported by Main Operating Bases that are required to provide backup support to them. A FOB also improves reaction time to local areas as opposed to having all troops on the main operating base.
- **Definition MOB** - A MOB is a term used by the United States military defined as a permanently manned, well protected base, used to support permanently deployed forces, and with robust sea and/or air access.
- **Definition COP** - A combat outpost is a detachment of troops stationed at a distance from the main force or formation, usually at a station in a remote or sparsely populated location, positioned to stand guard against unauthorized intrusions and surprise attacks; the station is occupied by troops, it is usually a small military base or settlement in an outlying frontier, limit, political boundary or in a foreign country.
- **Definition Firebase** - A temporary military encampment to provide artillery fire support to infantry operating in areas beyond the normal range of fire support from own base camps.
- **Definition Camp** - A semi-permanent facility for the lodging of an army. Camps are erected when a military force travels away from a major installation or fort during training or operations, and often have the form of large campsites.
- **Definition Base** - A facility directly owned and operated by or for the military or one of its branches that shelters military equipment and personnel, and facilitates training and

operations. In general, a military base provides accommodations for one or more units, but it may also be used as a command center, a training ground, or a proving ground. In most cases, a military base relies on some outside help in order to operate. However, certain complex bases are able to endure by themselves for long periods because they are able to provide food, water and other life support necessities for their inhabitants while under siege.

All locations are taken from Wikimedia's GeoHack program if available. We do not consider Firebases and COPs, which are smaller and often temporary outposts. In addition, we found or updated the information for the following cases:

1. COP/FOB Zangabad has been coded as FOB Pasab. This was the most likely location for a forward operating base close the Zhari/Panjwayi district border. Exact location determined as such using Wikimapia satellite imagery. It is coded as being in the district of Panjwayi.
2. Camp/FOB Hadrian location determined using Wikimapia satellite imagery.
3. Camp Russell location determined using Wikimapia satellite imagery in relation to Camp Holland.
4. Camp Arena, Camp Vianini, and Camp Stone are each in roughly the same area. Using Wikimapia imagery, we assume that Camp Arena, the only camp with an Italian Air Force presence, is located at the airfield in Hirat. Camp Vianini and Camp Stone were assigned their locations using Wikimapia imagery as well. We believe Camp Vianini to be at the location we chose based on the fact that an Italian artillery regiment was attacked at that location and we believe the Italian Army was the only major force at Camp Vianini. Camp Stone, which has multiple country forces at its location, is expected to be south of the airport and Camp Arena, according to Wikimapia data.
5. Camp Blackhorse determined using Wikimapia and various sources citing the camp to be adjacent to the Pul-e-Charkhi ANA compound.
6. Camp Clark determined using Wikimapia satellite imagery.

7. Camp Warehouse determined using Wikimapia satellite imagery.
8. Camp Phoenix location determined using google maps and Wikimapia satellite data.
9. Camp Invicta located using Wikimapia satellite data.
10. FOB Hamidullah located using Wikimapia satellite data. In Wikimapia, the location is described as FOB Nalay, the previous name of the base.
11. Camp Blessing located using Wikimapia satellite data.
12. FOB Joyce located using satellite data and with news articles stating that FOB Joyce is within/very close to the village of Serkanay.
13. Camp Wright located using Wikimapia and Google Maps satellite data; it is listed as 'USA Army Base' on the Wikimapia site.

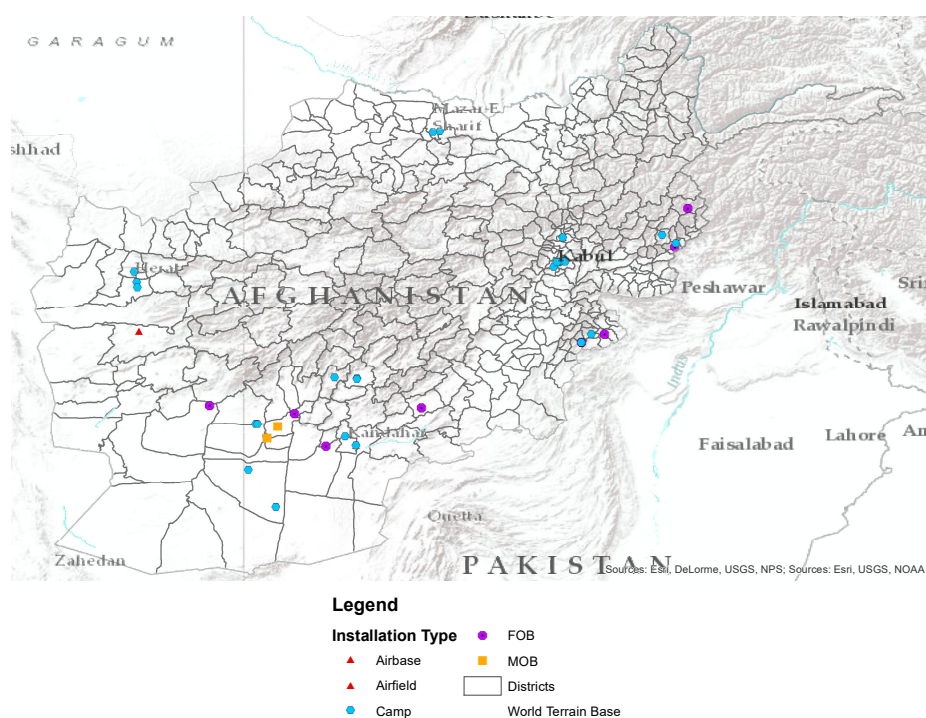


Figure H.6: Final Map of Located Military Installations

This map shows the geographic location of the bases that we identified. Some bases are not visible in this view as a result of closely overlapping with other bases, in which case the map displays only one symbol.

References

- Barrett, Christopher, & Christian, Paul. 2017. *Revisiting the Effect of Food Aid on Conflict: A Methodological Caution*. World Bank Policy Research Working Paper No. WPS 8171. 46
- Benmelech, Efraim, Berrebi, Claude, & Klor, Esteban F. 2012. Economic Conditions and the Quality of Suicide Terrorism. *The Journal of Politics*, **74**(1), 113–128. 37
- Berman, Eli, Shapiro, Jacob N, & Felter, Joseph H. 2011a. Can Hearts and Minds Be Bought? The Economics of Counterinsurgency in Iraq. *Journal of Political Economy*, **119**(4), 766–819. 47
- Berman, Eli, Felter, Joseph H, Shapiro, Jacob N, & Callen, Michael. 2011b. Do Working Men Rebel? *Journal of Conflict Resolution*, **55**(4), 496–528. 48
- Berman, Nicolas, & Couttenier, Mathieu. 2015. External Shocks, Internal Shots: The Geography of Civil Conflicts. *Review of Economics and Statistics*, **97**(4), 758–776. 3, 42
- Blair, Christopher W., & Wright, Austin L. 2021. Refugee Return and Conflict: Evidence from a Natural Experiment. *Becker Friedman Institute for Economics Working Paper Series*. 38, 39
- Bluhm, Richard, Gassebner, Martin, Langlotz, Sarah, & Schaudt, Paul. 2021. Fueling Conflict? (De) Escalation and Bilateral Aid. *Journal of Applied Econometrics*, **36**(2), 244–261. 42
- Bueno de Mesquita, Ethan. 2013. Rebel Tactics. *Journal of Political Economy*, **121**(2), 323–357. 37
- Cameron, A Colin, & Miller, Douglas L. 2015. A Practitioners Guide to Cluster-Robust Inference. *Journal of Human Resources*, **50**(2), 317–372. 44
- Cameron, A Colin, Gelbach, Jonah B, & Miller, Douglas L. 2011. Robust Inference with Multiway Clustering. *Journal of Business & Economic Statistics*, **29**(2), 238–249. 44
- Caulkins, Jonathan Paul, Kleiman, Mark, & Kulick, Jonathan D. 2010. *Drug Production and Trafficking, Counterdrug Policies, and Security and Governance in Afghanistan*. New York University Center on International Cooperation. 41
- Compton, Wilson, Jones, Christopher, & Baldwin, Grant. 2016. Relationship between Non-medical Prescription-Opioid Use and Heroin Use. *New England Journal of Medicine*, **374**(2), 154–163. 32
- Croicu, Mihai, & Sundberg, Ralph. 2015. UCDP Georeferenced Event Dataset Codebook Version 4.0. *Journal of Peace Research*, **50**(4), 523–532. 3, 4
- CSO. 2005, 2007/08, 2011/12. *National Risk and Vulnerability Assessment*. Kabul, CSO. 4, 5, 6, 8, 12
- CSO, & The World Bank. 2011. *Setting the Official Poverty Line for Afghanistan*. Tech. rept. 4
- Dart, R. C., Surratt, H. L., Cicero, T. J., Parrino, M. W., Severtson, S. G., Bucher-Bartelson, B., & Green, J. L. 2015. Trends in Opioid Analgesic Abuse and Mortality in the United States. *New England Journal of Medicine*, **372**(3), 241–248. 8
- Dell, Melissa, & Querubin, Pablo. 2018. Nation Building Through Foreign Intervention: Evidence From Discontinuities in Military Strategies. *The Quarterly Journal of Economics*, **133**(2), 701–764. 47
- Donaldson, Dave, & Hornbeck, Richard. 2016. Railroads and American Economic Growth: A Market Access Approach. *The Quarterly Journal of Economics*, **131**(2), 799–858. 9
- Dorronsoro, Gilles. 2005. *Revolution Unending: Afghanistan, 1979 to the Present*. Columbia University Press. 10, 70, 71

- EMCDDA. 2016. *European Drug Report - Trends and Developments*. Luxembourg: Publications Office of the European Union, 2016. 6
- FAO GAEZ, FAO/IIASA. Global Agro-ecological Zones (GAEZ v3.0). *Food and Agriculture Organization of the United Nations*. 68
- Giustozzi, Antonio. 2009. *Empires of Mud: The Neo-Taliban Insurgency in Afghanistan 2002-2007*. Columbia University Press. 10, 47
- Gordon, Raymond G. 2005. *Ethnologue*. SIL International. 6
- Harari, Mariaflavia, & La Ferrara, Eliana. 2018. Conflict, Climate and Cells: A Disaggregated Analysis. *Review of Economics and Statistics*, forthcoming. 13
- Henderson, J Vernon, Storeygard, Adam, & Weil, David N. 2012. Measuring Economic Growth from Outer Space. *American Economic Review*, **102**(2), 994–1028. 9
- Hijmans, Robert J, Guarino, Luigi, Cruz, Mariana, & Rojas, Edwin. 2001. Computer Tools for Spatial Analysis of Plant Genetic Resources Data: 1. DIVA-GIS. *Plant Genetic Resources Newsletter*, **127**, 15–19. 10
- IMF. 2005-2017. *International Monetary Fund - Database*. available at <http://www.imf.org/external/np/res/commod/index.aspx>. 13
- Kenan, Kristen, & Mack, Karin. 2012. Trends in Prescriptions for Oxycodone and Other Commonly Used Opioids in the United States, 2000-2010. *Open Medicine*, **6**(2), e41–e47. 8
- Kienberger, Stefan, Spiekermann, Raphael, Tiede, Dirk, Zeiler, Irmgard, & Bussink, Coen. 2017. Spatial Risk Assessment of Opium Poppy Cultivation in Afghanistan: Integrating Environmental and Socio-Economic Drivers. *International Journal of Digital Earth*, **10**(7), 719–736. 11, 68
- Manchikanti, Laxmaiah, II, Standiford Helm, Fellows, Bert, Janata, Jeffrey W., Pampati, Vidyasagar, Grider, Jay S., & Boswell, Mark V. 2012. Opioid Epidemic in the United States. *Pain Physician*, **15**(3 Suppl.), 2150–1149. 8
- Nunn, Nathan, & Puga, Diego. 2012. Ruggedness: The Blessing of Bad Geography in Africa. *Review of Economics and Statistics*, **94**(1), 20–36. 12
- OECD. 2016. *Organization for Economic Cooperation and Development - Consumer Price Index*. 4
- Olea, José Luis Montiel, & Pflueger, Carolin. 2013. A Robust Test for Weak Instruments. *Journal of Business & Economic Statistics*, **31**(3), 358–369. 32
- Shapiro, Jacob N. 2012. *Terrorist Decision-making: Insights from Economics and Political Science*. 5 edn. Vol. 6. 37
- Shapiro, Jacob N. 2013. *The Terrorist's Dilemma: Managing Violent Covert Organizations*. Princeton University Press. 37
- Shaver, Andrew C., & Wright, Austin L. 2016. Are Modern Insurgencies Predictable? New Evidence from the Afghanistan and Iraq Wars. *Working Paper*. 12
- Sonin, Konstantin, & Wright, Austin L. 2023. Rebel Capacity, Intelligence Gathering, and Combat Tactics. *American Journal of Political Science*. 40
- Staiger, Douglas O, & Stock, James H. 1997. Instrumental Variables Regression With Weak Instruments. *Econometrica*, **65**(3), 557–586. 32
- Sundberg, Ralph, & Melander, Erik. 2013. Introducing the UCDP Georeferenced Event Datas. *Journal of Peace Research*, **50**(4), 523–532. 3, 4
- Trebbi, Francesco, Weese, Eric, Wright, Austin, & Shaver, Andrew. 2017. Insurgent Learning. *NBER Working Paper No. w23475*, jun. 38
- UNDCP. 2000. *Annual Opium Poppy Survey*. 10

- UNODC. 2005-2016. Afghanistan Opium Survey: Socio-Economic Analysis. *Ministry of Counter Narcotics, Afghanistan and United Nations Office for Drugs and Crime, Vienna*. 3, 7, 9, 10
- UNODC. 2006-2013. *Afghanistan Opium Price Monitoring*. 8
- Van Hoolst, Roel, Eerens, Herman, Haesen, Dominique, Royer, Antoine, Bydekerke, Lieven, Rojas, Oscar, Li, Yanyun, & Racionzer, Paul. 2016. FAO's AVHRR-based Agricultural Stress Index System (ASIS) for Global Drought Monitoring. *International Journal of Remote Sensing*, **37**(2), 418–439. 12, 13
- Vanden Eynde, Oliver. 2018. Targets of Violence: Evidence From India's Naxalite Conflict. *The Economic Journal*, **128**(609), 887–916. 38
- Vanden Eynde, Oliver, Wright, Austin L., & Fetzer, T. R. 2017. *Security Transitions*. Mimeo. 39
- Webb, Matthew D. 2013. *Reworking Wild Bootstrap Based Inference for Clustered Errors*. Queen's Economics Department Working Paper No. 1315. 44
- Weidmann, Nils B, Rød, Jan Ketil, & Cederman, Lars-Erik. 2010. Representing Ethnic Groups in Space: A New Dataset. *Journal of Peace Research*, **47**(4), 491–499. 6, 7, 42
- Wiesmann, Doris, Bassett, Lucy, Benson, Todd, & Hoddinott, John. 2009. *Validation of the World Food Programme's Food Consumption Score and Alternative Indicators of Household Food Security*. IFPRI Discussion Paper No. 00870. 4, 5
- World Bank. 2005-2016. *World Development Indicators*. 4, 8
- Wright, Austin L. 2018. Economic Shocks and Rebel Tactics. *Working Paper*. 37, 38