ESSAYS IN APPLIED SPATIAL MICROECONOMICS

by

JAMES HERNDON

AMANDA ROSS, COMMITTEE CHAIR
ARTHUR ALLAWAY
TRAVISS CASSIDY
ERIK JOHNSON
JUNSOO LEE

A DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy
in the Department of Economics, Finance, and Legal Studies
in the Graduate School of
The University of Alabama

TUSCALOOSA, ALABAMA

2020

ABSTRACT

In the first essay, I examine the price behavior of consumer goods in the strategically vital country of Pakistan. Results show that prices converge both temporally and spatially. A wage-adjusted Consumer Price Index shows that Pakistani cities have converging costs of living. Finally, a novel measure of cointegration ranks the most and least economically integrated cities. Divergence does not occur along provincial, linguistic, or ethnic boundaries.

In the second essay, paper I examine private sector job growth in cities across the United States from 1990 to 2018. Defining "concentration" as a city's sectoral Herfindahl-Hirschman Index, I find that cities with greater economic concentration subsequently experience more job growth than comparable cities with less concentration. However, the skewed distribution of job growth by sector means that cities face a trade-off between risk and reward analogous to an investment portfolio.

In the third and final essay, we examine how changes in rainfall affect the persistence of conflict in Africa using fine-grained grid cell level data. Using Markov transition matrices, we examine the persistence of conflict in grid cells across the African continent and the likelihood of transitioning into and out of conflict. We incorporate the Markov probabilities into a panel logit model to analyze how monthly variations in rainfall affect the probability that an area transitions from peace to conflict. We find that peace is highly persistent across Africa, while violence is more transient. We also find that insufficient rainfall early in the wet season is associated with conflict in several regions.

DEDICATION

This dissertation is dedicated to Bruce Douglas Caples.

LIST OF ABBREVIATIONS AND SYMBOLS

α	Intercept Term for a Regression
β	Coefficient Term for a Regression
ϵ	A stochastic error term
μ	Mean of Prices
σ	Standard Deviation of Prices
ACLED	Armed Conflict Location and Event Database
BLS	Bureau of Labor Statistics
CHIRPS	Climate Hazards Group Infrared Precipitation with Station
COV	Coefficient of Variation
CPI	Consumer Price Index
FIPS	Federal Information Processing System
FIRE	Finance, Insurance and Real Estate
FOD	Forward Orthogonal Deviation
HHI	Herfindahl-Hirschman Index
MSA	Metropolitan Statistical Area
NAICS	North American Industry Classification System
PBS	Pakistan Bureau of Statistics
QCEW	Quarterly Census of Employment and Wages
rv	Relative Volatility
SIC	Standard Industrial Classification
SPI	Sensitive Price Indicator
WPI	Wholesale Price Index
$\mathbf{X}$	A vector of control variables

ACKNOWLEDGMENTS

I was fortunate to work with and learn from faculty in the Department of Finance, Economics and Legal Studies. In particular, I am grateful for the time and effort that my committee members put in to making each essay far better than anything I could have produced alone. In particular, I should thank Dr. Amanda Ross for offering me the opportunity to present my work at the Regional Science Association International conference in 2018. I would also like to express my gratitude to Drs. Matt Holt and Greg Givens.

I am also grateful for the friendship and camaraderie I found with my fellow PhD students. In particular, I was lucky to work alongside Shamar Stewart, Matthew Hampton, Mia Goodnature, Hasan Isomitdinov, and Olivia Lu.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

A NOTE ON PRICE BEHAVIOR IN URBAN PAKISTAN

1.1 Introduction

President Obama once called the disintegration of Pakistan "his biggest single national security concern" Sanger (2012). This article argues that such a scenario is unlikely to arise for strictly economic reasons as Pakistani cities provide effective and efficient markets for consumer goods, with firms and consumers buying and selling across provincial and linguistic barriers, and a lack of significant differences in the cost of living in different cities.

Price deviations have been argued to convey information about market integration or lack thereof. Engel and Rogers (1996) find that the international border between United States and Canadian cities contributed more to relative price volatility than thousands of miles of distance within each country. This occurred despite a free trade agreement, a common language, and a similar legal system. Subsequent articles have focused more on prices within a single nation to eliminate confounding factors like borders and exchange rates from the analysis, such as Parsley and Wei (1996) in the United States, Fan and Wei (2006) in China, and Morshed et al. (2006) in India.

1.2 Pakistan's Urban Context

The British carved Pakistan out of the larger India colony in 1947, imposing a country on diverse populations they had ruled more like a "military fiefdom" than a proper Crown Colony, per Kaplan (2012). Pakistan's first rulers thus inherited a limited state; today the central government still does not control sixty percent of its nominal territory, per Rumi

(2016).

The four provinces of Pakistan roughly correspond to four major ethnic and linguistic groups: Punjab, Sindh, Khyber Pakhtunkhwa and Baluchistan. The rural-urban divide within provinces also matters. Blank et al. (2014) found what while the residents of major cities in Pakistan enjoyed better access to education and healthcare, they increasingly saw their economic situation as worse than those living in deprived rural areas. The benefits of migration often fail to meet their expectations, and that dissatisfaction might grow much more dangerous if consumers suddenly faced rising prices.

1.3 Model and Econometric Techniques

Like Engel and Rogers (1996), I assume vendors sell goods comprised of two intermediate inputs: one tradeable and one nontradeable. The nontradable component includes transportation costs that prevent prices from fully converging. Still, if prices vary too much arbitrage can occur. Since transportation and opportunity costs both increase with distance, even in a well-integrated market the relative price of a good in locations j and k should fluctuate within a range proportional to the distance between j and k. If the relative price of a good in two places wanders within these bounds, prices appear to behave like a unit root while remaining within the bounds of the model. We should not expect to see perfect convergence across all locations or even between two cities in close proximity. However, we can test whether the price for a particular good in a particular city ever diverges from the national market, or if convergence occurs spatially within each province to the exclusion of cities in other provinces. Either outcome would indicate a market failure.

To identify goods that could indicate the presence or absence of such market failure, this article uses each good's coefficient of variation (COV) as a criteria for inclusion. If across all cities the ratio of a good's standard deviation, σ , to its mean, μ, over all time periods, t, equals zero, then there is no opportunity for arbitrage. Every good tested in this article had a nonzero COV. We expect the COV, however, to differ by item, because higher transportation costs for easily damaged goods means their relative prices could vary more before offering an opportunity for profit. Perishable goods' prices might also vary more, as they require either expensive storage or immediate resale.

I examine price behavior at the level of individual consumer goods with two tests: one for each good temporally without city effects, and another that measures each good's convergence spatially between cities. First I transformed prices using the forward orthogonal deviation (FOD) introduced by Arellano and Bover (1995) to remove individual city effects. Hayakawa (2009) showed that FOD, which subtracts the average of all future values, outperforms the first difference method in both estimator bias and root-mean-squared error. I represent this transformed price for a given time period t as *price**. Regression of a price on its first lag, shown in Equation 1, conveys how quickly a price reverts to its mean. I follow Arellano and Bover (1995) in using the second lag in levels as an instrumental variable for the lagged price.

$$price^*_{i,t} = \alpha + \beta price^*_{i,t-1} + \epsilon_{i,t} \tag{1.1}$$

Phillips and Sul (2007) devised a test for panel data with the dynamics of spatial convergence in mind. Phillips and Sul's log t convergence tests the null hypothesis that each city's divergence from a common trend tends toward zero. If $x_{i,t}$ represents the price of a good in city i at time t, let $h_{i,t}$ represent the price of each good in city i in time i, normalized to 1, as shown in Equation 2. Equation 3 defines H_t as a scale invariant measure of that good's price dispersion during period t. Converging prices cause H_t to fall, which raises the left-hand side of the regression shown in Equation 4.

$$h_{i,t} \equiv \frac{X_{i,t}}{(1/N)\sum_{i=1}^{N} X_{i,t}} \tag{1.2}$$

$$H_{i,t} \equiv (1/N)\sum_{i=1}^{N}(h_{i,t}-1)^2 \tag{1.3}$$

$$\ln\frac{H_1}{H_t} - 2\ln(\ln t) = \alpha + \beta\ln t + \epsilon_t \tag{1.4}$$

Phillips and Sul (2007) assumed a time-varying factor model for prices that includes a trend and an idiosyncratic component. They derive the log t test from the semi-parametric equation used to describe the model's idiosyncratic component. Under these conditions, a t-test of the β-coefficient in Equation 4 tests for convergence.

Phillips and Sul (2007) calculated Consumer Price Indices (CPI) for 19 cities in the United States from 1960 to 2000. After rejecting overall convergence, Phillips and Sul tested for "club convergence." They ordered the indices by last observation, ran the log t regression for the two most expensive cities, then kept adding more cities one at a time until the t-statistic fell below the 5 percent threshold. Whatever combination maximized the t-statistic they designated as the "core group" They then added cities to the core group one at a time, leaving cities in so long as the t-statistic exceeds zero, the threshold that Phillips and Sul found minimized the sum of Types I and II errors. Whatever combination of cities met this cutoff was the first "club." Then they repeated the process with the remaining cities. A city that does not fall into any club Phillips and Sul labeled "divergent." The log t convergence results in this study were obtained using the Stata module introduced by Du (2017).

Applying this to Pakistan, I constructed a monthly CPI that controls for wages, a rough measure of each city's cost of living. If I subject those wage adjusted indices to the log t test, I would not expect to see any divergent cities. Given free labor movement, a city with high cost of living would either need to provide amenities valued by its citizens or risk losing inhabitants to migration, whereas a low cost of living city would attract them. "Clubs" inside of Pakistan are no more evidence of market failure than those in the United States, assuming that they do not emerge along provincial or ethnic lines.

Finally, if the Law of One Price held perfectly, then any differences I encounter should reflect either differences in markups or transportation costs. For example, according to the World Bank (2016) 95 percent of Pakistan's international trade goes through the southern port city of Karachi . So the price of an imported good in every other city reflects the cost of the good in Karachi plus the local markup and transportation costs. If the price in Karachi rises by 100 rupees, then the cost in other cities should rise by a predictable amount and the relative price of a good in two cities will not vary over time. Alternatively, an unstable relative price indicates that two cities are not integrated. Defined in Equation 5, the "relative volatility" of good i in cities j and k is the standard deviation of the normalize price difference across all time periods t.

$$rv_{j,k} \equiv sd\left[\frac{p_{j,t} - p_{k,t}}{(p_{j,t} + p_{k,t})/2}\right] \tag{1.5}$$

If one region had significantly higher scores across most or all of its cities, I may infer that something about that region inhibits arbitrage.

1.4 Data

Every Thursday, the Pakistan Bureau of Statistics (PBS) visits multiple markets to record retail prices for 53 "essential commodities" in 17 cities across Pakistan: the capital, Islamabad, 8 cities in Punjab (Rawalpindi, Gujranwala, Sialkot, Lahore, Faisalabad, Sargodha, Multan, and Bahawalpur), 4 in Sindh (Karachi, Hyderabad, Sukkur, and Larkana), 2 in Khyber Pakhtunkhwa (Peshawar and Bannu), and 2 in Baluchistan (Quetta and Khuzdar), per PBS Methods (2018). Field offices contact shopkeepers directly in order to avoid nominal price controls. A weighted average of these prices forms the "Sensitive Price Indicator" (SPI). Prices for individual goods in different cities have been posted since January 2012, per PBS Indices (2018). This article uses 84 monthly observations from January 2012 to December 2018.

The PBS also collects prices for a broader CPI (487 items) and a Wholesale Price Index (463 items). Prices for every good in the wider CPI and Wholesale Price Index are not publicly available. For the time period covered in this article, we see an early surge in inflation followed by a gradual and irregular decline, per the State Bank of Pakistan (2018).

Almost every week, the PBS posts daily wages in rupees for five occupations in the same cities as the SPI. Three trades have uninterrupted observations for every month: carpenter, electrician, and mason. The wages used in this article are the first available wage data for each month from January 2013 to December 2018. The "wage" of a particular city refers to the mean of these three daily wage rates, shown normalized to 100 in Table 1.2, below. Note that when we control for wages we lose the 2012 price data. Also, variation in local income levels could cause price variation via demand pull inflation. I do not believe that occurred because the cities' wages displayed a mean-reverting COV. I also ran Phillips and Sul's convergence test on wages and no city diverged.

1.5 Results

I exclude three goods with an average COV equal to zero. See Table 1.1, column 2. Utilities such as electricity charges and telephone calls are not subject to arbitrage, while the manufactured name-brand "Bata" sandals are uniform across the country. Gas charges are excluded due to missing observations in Bannu and Khuzdar. Near-uniform prices do not imply integration or even a functioning market. For instance, sugar's COV was 0.033. Salman (2015) explained how price support for sugar cane resulted in massive subsidies for sugar cane growers. Bananas had one of the highest COV: 0.294. That confirms the intuition that perishable, easily damaged goods display more variable prices.

The closer the β-coefficient on the first lag to 1, the more "sticky" the price. Seasonal and perishable items vary more month-to-month as the supply is irregular, leaving them with lower coefficients. For example, in column 3 of Table 1.1, consider onions and tomatoes (items 23 and 24). Beef and mutton (items 6 and 7) displayed a near-exponential increase in prices in this period, much to the consternation of Pakistanis. Nonperishable goods that are easy to transport and store should have more stable prices and coefficients closer to one, as seen in clothing (items 35–38). I also observed stable prices among packaged goods (items

11 and 13–15) and hotel prepared meals subject to "menu costs" that inhibit frequent price changes (items 31–33).

When examining goods spatially by finding their Phillips and Sul log t convergence clubs, I obtain the results shown in Table 3, column 5. Out of 833 time series (17 cities by 49 items) only 58 diverged from any club, or about 6.9%. The two cities most likely to diverge were Sargodha (Punjab) and Hyderabad (Sindh) with seven each. Whatever their political grievances, citizens in Khyber Pakhtunkhwa and Baluchistan enjoy a market for consumer goods that functions as well as it does in the rest of the country. When I depart from the specification in Phillips and Sul (2007) and exclude only the first 15% of observations the number of divergent city prices rises to 81 (9.7%)

The average good had two or three clubs. In their article using the Phillips and Sul test on a wide range of goods, Dang et al. (2018) looked at 85 tradable prices across 15 Asian cities from 1990 to 2008 and found that only 5% converged into a single club. While this study is not directly comparable to Dang et al. (2018), it does highlight the integration of the markets for five of the goods that converged into a single club: wheat, live chickens, bananas, gram (chickpeas), and hotel-served beef.

To understand how this price behavior affects quality of life, I needed a measure of price levels overall for each city. An official CPI or SPI for different Pakistani cities is not available, so I constructed one using the 49 goods used in this article. The remaining items were weighted by modifying the full Pakistani CPI weights published by the State Bank of Pakistan (2018). Certain categories were not counted at all, such as health, transport, and education. The goods in the index fell into categories that total 0.81, so I adjusted food and nonalcoholic beverages, for example, from 0.348 to 0.430 (i.e., 0.348/0.81). Goods within each category received equal weight, so all 30 items in the food and nonalcoholic beverage category were weighted 0.0143 (0.430/30).

To construct each city's CPI, I began by taking the product of every price and its weight, then summing for that month. Then I divide this by "wage" for the cost of living. Recall that I defined "wages" as the mean daily wage for a carpenter, an electrician, and a mason in that city month. When I subjected the 17 cities' wage-adjusted cost of living time series to the Phillips and Sul test, all cities converged into one group. I also found that the cost of living before controlling for wages displayed less variation than the wage-controlled version. This is consistent with the expectation that goods are more mobile than labor.

This lack of a single divergent city should not be surprising since Memon (2005) classified between 1.7% and 4.7% of Pakistan's population as economic migrants. In spite of regional tensions, language differences, and other obstacles, the freedom of movement ensures living standards remain somewhat even across the country.

I calculated each city's "relative volatility" by constructing a 49 17 matrix where each entry is the solution to Equation 5 for one item and city. The last column of Table 1.2 represents the sum of all entries in each city's column, normalized to a mean of 100. Islamabad is the most integrated city in Pakistan according to Equation 5. After Islamabad, the next four most integrated cities are all in Punjab, and thus in close proximity both spatially and linguistically. These results are summarized in Table 1.2, below, which also lists the population for each city according to PBS Census (2017).

Baluchistan's cities are the least integrated by far. So while my results support the notion of Pakistan as an integrated market, distance does inhibit integration somewhat.

To understand what drives relative volatility, I used the normalized relative volatility for all city pairs as a dependent variable in a cross-sectional ordinary least squares regression. Results appear in Table 1.3. After controlling for individual city effects, a provincial border between two cities has the same effect as roughly 589 km of distance. This result is an order of magnitude lower than the estimate for the border between the United States and Canada in Engel and Rogers (1996).

1.6 Conclusion

This article used price behavior as a proxy for Pakistan's economic integration. For consumer goods in a free market, prices in urban Pakistan behave according to basic economic theory: a variety of local conditions prevent perfect convergence, but across all goods prices do not diverge either temporally or spatially.

These results suggest that, despite other problems, Pakistan today is an integrated market with prices robust to many of the shocks that upset less stable regimes.

Table 1.1: Sensitive Price Indicator Goods

Item	COV	β	Half-Life	Clubs	Divergent Cities	CPI Weight
Wheat	0.052	0.854	4.38	1	0	0.014
Wheat Flour, Bag	0.056	0.888	5.813	2	3	0.014
Rice Basmati Broken, (AQ)	0.143	0.971	23.671	4	0	0.014
Rice Irri-6 (Punjab/Sindh)	0.149	0.962	17.681	1	2	0.014
Bread Plain, Medium Size	0.09	0.969	22.158	2	1	0.014
Beef With Bone, (AQ)	0.083	1.005	.	2	1	0.014
Mutton, Average Quality	0.097	1.007	.	2	3	0.014
Chicken Farm, Broiler, Live	0.116	0.352	0.664	1	0	0.014
Milk, Fresh, Unboiled	0.112	0.98	33.688	2	0	0.014
Curd (Dahi)	0.14	0.987	51.215	3	0	0.014
Powdered Milk, Nido, Polybag	0.014	0.964	18.862	4	5	0.014
Eggs Hen, Farm	0.066	0.543	1.136	4	0	0.014
Mustard Oil, Average Quality	0.146	0.988	57.147	2	0	0.014
Cooking Oil, Tin, (SN)	0.02	0.963	18.568	3	1	0.014
Vegetable Ghee, Tin, (SN)	0.02	0.963	18.177	2	0	0.014

Continued on next page

Table 1.1 – *Continued from previous page*

Item	**COV**	β	**Half-Life**	**Clubs**	**Divergent Cities**	**CPI Weight**
Vegetable Ghee (Loose)	0.071	0.962	17.996	5	0	0.014
Bananas	0.294	0.704	1.972	1	0	0.014
Pulse Masoor, Washed	0.109	0.961	17.39	6	1	0.014
Pulse Moong, Washed	0.093	0.957	15.641	6	2	0.014
Pulse Mash, Washed	0.112	0.968	21.526	3	5	0.014
Pulse Gram, Washed	0.092	0.946	12.472	1	0	0.014
Potatoes	0.218	0.769	2.637	3	2	0.014
Onions	0.184	0.646	1.588	2	4	0.014
Tomatoes	0.19	0.36	0.678	5	0	0.014
Sugar, Refined	0.033	0.886	5.732	3	1	0.014
Gur, Average Quality	0.129	0.881	5.459	4	4	0.014
Salt Powder, Loose, Lahori	0.126	0.99	69.677	4	2	0.014
Red Chilly Powder, Loose	0.129	0.939	10.987	2	0	0.014
Garlic	0.152	0.922	8.523	2	0	0.014
Tea, Lipton Yellow Label, Packet	0.02	0.958	16.224	4	2	0.014
Cooked Beef, Average Hotel	0.21	0.991	80.18	1	0	0.005
Cooked Daal, Average Hotel	0.214	0.992	81.316	2	0	0.005

Continued on next page

Table 1.1 – *Continued from previous page*

Item	COV	β	Half-Life	Clubs	Divergent Cities	CPI Weight
Tea Prepared, Average Hotel	0.125	0.984	43.616	1	3	0.005
Cigarettes, K-2, 20's	0.048	0.964	18.788	3	0	0.017
Long Cloth	0.226	0.973	25.561	3	0	0.016
Shirting	0.238	0.991	74.275	2	2	0.016
Lawn	0.288	0.976	28.585	5	2	0.016
Georgette	0.155	0.977	29.812	3	0	0.016
Sandal Bata (Gents)	0	.	.	.	.	.
Chappal Spounge Bata (Gents)	0.001	0.642	1.562	1	0	0.016
Sandal Bata (Ladies)	0.008	0.951	13.898	1	1	0.016
Elec. Charges, Upto 50 Units	0	.	.	.	.	.
Gas Charges, Up to 3.3719 MMBTU	0.09	.	.	.	.	.
Kerosene Oil	0.103	0.972	24.14	4	0	0.061
Firewood Whole	0.2	0.99	66.878	3	1	0.061
Energy Saver (14 watts)	0.061	0.96	17.038	3	2	0.052
Washing Soap, (Nylon), (SN)	0.295	0.987	51.392	3	1	0.014
Match Box, Regular	0.227	0.981	35.999	1	1	0.061
Petrol, Super	0.012	0.952	14.075	1	1	0.061

Continued on next page

Table 1.1 – *Continued from previous page*

Item	**COV**	β	**Half-Life**	**Clubs**	**Divergent Cities**	**CPI Weight**
Hi Speed Diesel (HSD)	0.011	0.957	15.824	1	1	0.061
LPG Cylinder, 11kg	0.073	0.831	3.751	2	0	0.061
Telephone Local Call	0	.	.	.	.	.
Soap, LifeBuoy	0.014	0.954	14.566	4	4	0.014

Table 1.2: City Summary

City	Population	Average CPI	Average Wage	Relative Volatility
Islamabad	1,014,825	109.52	124.49	85.37
Rawalpindi	2,875,516	106.02	123.35	86.73
Gujranwala	2,948,936	99.89	98.24	98.25
Sialkot	1,143,362	100.64	97.1	95.2
Lahore	11,126,285	103.07	96.86	96.41
Faisalabad	3,760,328	99.96	86.73	100.84
Sargodha	1,091,045	96.84	94.43	99.37
Multan	2,058,290	98.63	93.61	91.35
Bahawalpur	1,171,258	99.51	77.55	98.06
Karachi	14,910,352	98.53	107.18	99.52
Hyderabad	1,832,755	95.74	99.1	97.19
Sukkur	720,115	94.79	81.3	99.65
Larkana	701,637	96.19	88.17	107.2
Peshawar	1,970,042	107.22	96.7	97.63
Bannu	49,965	94.71	113.34	100.08
Quetta	1,001,205	102.74	112.02	116.88
Khuzdar	277,136	96.01	109.84	130.26

Table 1.3: City-Pair Relative Volatility

	Dependent variable:	
	Relative Volatility	
	(1)	(2)
Distance (kilometers)	0.000531*	0.000612***
	(0.000299)	(0.000206)
Border Dummy	0.872***	0.361**
	(0.172)	(0.153)
Constant	−0.943***	2.340***
	(0.153)	(0.635)
Observations	136	136
R^2	0.317	0.855
City Controls	No	Yes
Note:	*p<0.1; **p<0.05; ***p<0.01	

CHAPTER 2

THE ROLE OF CONCENTRATION AND DIVERSITY IN THE GROWTH OF CITIES

2.1 Introduction

Mayor Mike Bloomberg's 2020 campaign website made a specific claim regarding his stewardship of New York City: "Mike focused job-creation efforts in industries that diversified the city's economy… in doing so, Mike and his team wrote one of the great comeback stories in American history" Bloomberg (2020). But why should a city's internal economic diversity matter at all?

In one sense, Bartik (1991) answered this question: cities grow when their local industries experience a positive shock nationally. Unfortunately, municipal leaders cannot predict future labor and technology shocks. So perhaps a city with more diverse firms could benefit from positive shocks while limiting its potential exposure to shrinking sectors. In this paper I use data from the Bureau of Labor Statistics (BLS) to show that cities do face that risk-reward trade-off. Furthermore, my reduced-form regressions quantify the relative contributions of diversity and concentration to job growth in American cities in recent decades. Economic concentration, as measured by the Herfindahl-Hirschman Index (HHI), lead to greater job creation in recent decades. In contrast, economic diversity, measured as the total number of sectors with at least one employee in that city, preceded lower job growth. These findings are robust to inclusion of confounders and the exclusion of outliers.

The remainder of this paper proceeds as follows. The "Background" section situates this work in the context of recent literature on agglomeration and job creation. The "Empirical Framework" outlines my estimation strategy while "Data and Descriptive Statistics" describes and summarizes the BLS's Quarterly Census of Employment and Wages (QCEW). The "Results and Discussion" section quantifies the role that concentration and diversity have played in recent years. The "Conclusion" offers implications for policy and future research.

2.2 Background

Much of the literature on agglomeration and city size focuses on productivity. For Rosenthal and Strange (2004) the terms are almost synonymous "...the forces are known variously as agglomeration economies or external economies of scale." Duranton and Puga (2004) argued that larger cities contribute to greater productivity via three non-exclusive mechanisms: sharing, matching, and learning. While a city's overall size matters, economic concentration may also play a substantial role: it is easier to share infrastructure, poach workers, or exchange ideas with similar firms. The theoretical model derived in Duranton and Puga (2004) assumes that some firms within a city produce a tradable final good while other firms produce intermediate goods that must be consumed by the exporting firms; in that model every city specializes in a single sector export in equilibrium. While an exaggeration, that result does offer an explanation for the results below, where more concentrated cities see faster growth. Ellison et al. (2010) provided empirical support for this notion by examining the collocation of industry-pairs such as cotton mills and thread mills. But while agglomeration and concentration may lead to innovation and rising output, the effect for employment is ambiguous. Depending on the elasticity of demand, a positive technology shock may lower employment in that industry faster than it can attract new firms. And while productivity is essential for long-run wealth creation, voters and politicians tend to focus more on jobs.

The literature on job creation tends to focus on the firm's characteristics, not the city's. Moscarini and Postel-Vinay (2012) found that large firms destroy proportionally more jobs during recessions than small firms. Adelino et al. (2017) found that younger firms account for most new jobs in the US. Sedláček and Sterk (2017) showed that firms founded during recessions tend to remain smaller than comparable firms founded during expansions, even after the economy recovers. A separate strand of research considered the role of wider economic diversity in job creation. For example, Glaeser et al. (1992) found that employment in a city-industry grew faster when the rest of that city was less specialized. Glaeser et al. (2015) showed that cities dominated by large firms have fewer entrepreneurs in subsequent generations. Clarke et al. (2016) explicitly tied urban agglomeration to job creation, but that paper considered employment growth at the firm level and did not consider any variables below the city level. Meanwhile, Hellerstein et al. (2015) studied labor markets at the neighborhood level. To the best of my knowledge, this is the first paper to measure the relationship between Metropolitan Statistical Area (MSA) level concentration/diversity and MSA-level employment growth.

The paper also contributes to the recent literature on economic heterogeneity within cities. Brunelle (2013) highlighted the increasing spatial concentration of work by function within industries, showing that senior management, scientists, and engineers all became more concentrated in major Canadian metro areas from 1971 to 2006. Markusen and Venables (2013) constructed a two-city model in which falling communication costs leads to the outcome described in Brunelle (2013). Labor force diversity also matters: Ottaviano and Peri (2006) found that cities with a higher concentration of foreign-born residents subsequently enjoyed faster increases in wages. Gabe and Abel (2012) found that the benefits of industrial concentration have increased for more specialized workers in the US. Diamond (2016) focused on workers sorting geographically by education, and how this may have contributed to inequality in recent decades.

2.3 Empirical Framework

Like much of the empirical literature on agglomeration, this paper relies on reduced form estimations. As Duranton and Puga (2014) noted: "...the most natural and direct way to quantify agglomeration economies is to estimate the elasticity of some measure of average productivity with some measure of local scale, such as employment density or total population." But in order to discern the relationship between job creation, concentration, and diversity, I instead make the percentage job growth for the entire city the dependent variable. All regressions in this paper control for population because larger cities typically display less concentration and greater diversity. See Figures 2.1 and 2.2, below. In order to remove the confounding effects of larger trends distinct from agglomeration at the city level I also controlled for the initial location quotients for jobs in the following sectors: retail, manufacturing, high tech, and finance, insurance and real estate (FIRE). The location quotient for an industry is the percentage of the workforce in that industry within a city, divided by the percentage of the workforce in that industry nationally. If 10 percent of a city's private sector employees works in a particular industry as opposed to 5 percent of the national workforce, then the location quotient is 10/5=2 . For a sense of possible ranges: the New York City MSA in 2018 has location quotients of 3.3 and 0 for North American Industry Classification System (NAICS) 523 (Securities, Commodity Contracts, and Other Financial Investments and Related Activities) and NAICS 113 (Forestry and Logging), respectively. Smaller cities can display a higher right skew: Tuscaloosa, Alabama's location quotient for NAICS 113 was 68. My baseline equation is shown below:

$$\%growth_{i,t} = \alpha + \beta concentration_{i,t-1} + \gamma \boldsymbol{X}_{i,t-1} + \epsilon_{i,t} \tag{2.1}$$

Where i represents a city, t a time period, $\boldsymbol{X}$ is a vector of the controls mentioned above, and ϵ is a stochastic error term. In this paper, I define industries by their NAICS code, in which more digits indicate a higher level of specificity. For example, NAICS codes 11, 111, and 1111 represent "Agriculture, Forestry, Fishing and Hunting," "Crop Production," and "Oilseed and Grain Farming," respectively. I define "concentration" in equation 2.1 as each MSA's Herfindahl Index at the 3-digit NAICS level. This index ranges between 0 and 1, where 1 represents total concentration. To find the HHI I find the sum of every industry's squared share of total employment. For example, if one half of a city works in one industry while the other half worked in another, then that city's concentration would be $(1/2)^2 + (1/2)^2 = 0.25$. Another city with greater concentration might have ratios of three-fourths and one-fourth, and thus concentration of $(3/4)^2 + (1/4)^2 = 0.625$.

I also ran a second set of regressions on initial diversity, defined as the number of unique 3-digit NAICS codes with at least one employee in that MSA-year. The argument for agglomeration from diversity is one of "cross-pollination" whereby innovations spread across industries. In theory this measure of diversity might have only a loose connection with concentration: two cities with the same number of distinct NAICS industries could still have very different levels of concentration. A negative coefficient on the diversity independent variable in equation 2.2, below, could imply that agglomeration's effects are stronger among more related industries.

$$\%growth_{i,t} = \alpha + \beta diversity_{i,t-1} + \gamma population_{i,t-1} + \epsilon_{i,t} \tag{2.2}$$

While this paper does not make explicit causal claims, by regressing growth on initial conditions I avoid the issue of reverse causality. My identification strategy rests on the assumption that other factors related to the changing number of jobs, such as weather and the quality of local government, are orthogonal to initial concentration. I also assume that any measurement errors in the number of jobs are randomly distributed across cities.

2.4 Data and Descriptive Statistics

In this paper I use annual QCEW data from 1990-2018, which the BLS derives from information submitted by all establishments subject to state unemployment insurance laws. Here I pause to note that in 2001, the BLS switched from the Standard Industrial Classification (SIC) job coding system to NAICS. The BLS recast their older data by NAICS only back to 1990. This data set covers roughly 95 percent of US employees and includes part-time and temporary workers BLS (2020). I define a city as the MSA identified by a unique Federal Information Processing System (FIPS) code. For the years 1990-2018, 353 cities have unbroken observations. I made one significant adjust to the data. The MSA "Los Angeles-Long Beach-Santa Ana, CA MSA" appeared in the years 1990-2012 under FIPS C3110. In 2013 the MSA "Los Angeles-Long Beach-Anaheim, CA MSA" appears under the FIPS code C3108. Given the significance of losing the second largest metropolitan area, I chose to treat this as one continuous observation. The BLS data provides annual numbers for average number of employees in specific industries at different levels of specification per the North American Industrial Classification System (NAICS). I define a city's "population" as the sum of all annual employment for that MSA across all private sector 4-digit NAICS codes for that year. For example, in 2018 Tuscaloosa had a total working "population" of 39,247. This count included employees in 83 unique occupations, the largest of which was 8,696 people working in Industry Code 7225, "Restaurants and Other Eating Places." I chose this measure of population because models of city size such as Duranton (2007) and Desmet and Rossi-Hansberg (2014) focus on firms and workers as opposed to government employees, schoolchildren and retirees.

Table 2.1, below, displays the relative size for MSAs along the distribution in the first and last years in the data. Cities along the entire distribution grew from 1990 to 2018. Table 2.2, below, describes how employment changed across all MSAs by decade. Note that while median MSA job growth in the 1990s was over 26 percent, private employment in the largest metro area actually fell during the 1990s. In addition, the overall rate of MSA employment growth dipped in the 2000s, likely as a result of capturing job growth over a decade bookended by recessions. Figure 2.1, below, plots the logarithm of HHI at the 3-digit NAICS codes for MSAs against the logarithm of total employment in 2018. As expected, larger cities tend to be less concentrated, although there is substantial variation that constitutes the main independent variable of interest.

City HHI grew more concentrated from 1900 to 2018 at both the 3-digit NAICS level;

the mean and median rose from 0.058 and 0.048 to 0.061 and 0.055, respectively. The Kolmogorov-Smirnov test confirmed that the 2018 distribution was stochastically greater than the 1990 with a high degree of confidence. This trend also occurred at the 4-digit NAICS level.

Dividing cities into two groups according to their initial concentration in 1990 provides suggestive evidence regarding the risks and rewards of concentration. Cities with higher HHI should benefit more if they are concentrated in an industry that sees rapid growth, but that also exposes them to downside risk. To put it bluntly: sometimes you're Silicon Valley, sometimes you're Detroit. Figure 2.3, below, captures this relationship graphically, with each city's HHI at the 3-digit level in 1990 on the horizontal axis and the percentage increase in private-sector jobs from 1990-2018 on the vertical axis. Cities in the bottom half of the HHI averaged 59 percent growth, with a standard deviation of 59 percentage points. The upper half averaged 99 percent growth with a standard deviation of 100 percentage points. The risk-reward trade-off is clear when I consider cities with the worst job creation records. For more concentrated cities, the ten slowest growing cities saw an average of -25 percent growth in jobs. For the worst performers in the less concentrated half, the average of the ten worst cities was -14 percent. This trend also holds for HHI at the 4-digit level.

Much of this risk results from the skewed the distribution of job growth across industries. 57 of the 89 3-digit NAICS industries grew over time, and a majority of the total increase occurred due to a few industries. To understand why, recall that most industries are relatively small relative to overall employment; the median number of employees in a 3-digit NAICS code in 1990 was about 440,000. The median rate of growth over the next 28 years was 18 percent. The industry that contributed the most was Food Services and Drinking Places (NAICS code 722), accounting for 17 percent of all MSA job growth from 1990 to 2018, followed by Ambulatory Health Care Services (NAICS code 621) with roughly 14 percent of new jobs. Over 80 percent of total job creation occurred in just ten industries. As for the losers, not even Mayor Bloomberg could have done much for a city specializing in Computer and Electronic Product Manufacturing (NAICS code 334). However promising that industry might have looked in 1990, jobs in that NAICS code fell from 1,507,691 in 1990 to 726,355 in 2018. That loss of 781,336 jobs was greater than the working population of all but 16 MSAs in 1990.

2.5 Results and Discussion

Initial concentration and diversity have displayed both statistically and economically significant impacts on MSA job creation since 1990. Table 2.3, below, shows the long-run effect of industrial concentration in 1990 on growth over the next 28 years. The overall distribution of cities grew more even in this time period, as shown by the negative and significant coefficient for the logarithm of population in 1990. As expected, the coefficients on both manufacturing's and services' location quotients are highly significant and have the expected sign. Despite these controls, the original degree of concentration still has an economically and statistically significant effect on growth. To move from the 25th to the 75th percentile, a city in 1990 would have needed its HHI to increase from 0.0390 to 0.0635. All else equal, that would have resulted in roughly 21 percent greater growth from 1990-2018. For perspective, the median total growth from 1990-2018 across 353 MSAs was roughly 56 percent.

In Table 2.4, below, I considered growth by decade. The role of concentration in subsequent job growth has been increasing in size and significance since the 1990s. This occurs alongside other economically significant trends: the decline in manufacturing appears to have peaked in the 2000s while the recent loss in high tech jobs reflects in part the loss of jobs in Computer and Electronic Product Manufacturing noted above.

Economic diversity in and of itself does not appear to lead to growth. Note that here I switched to the 4-digit NAICS code because it allows me to capture greater variation across cities. Figure 2.2, below, shows that MSAs of the same size can employ workers across a greater or lesser range of occupations at the 4-digit NAICS levels. If job growth is unpredictable, one might expect that cities with a more diverse work force are in position to benefit from more potential shocks. The evidence does not bear this out. After controlling for population, cities with jobs across more 4-digit NAICS industries display less job growth over time. Table 2.5, below, shows that having a few people in several fields at the start of the period is not enough. In fact, for a given city, having 25 more NAICS in 1990 is associated with 38 percent less growth by 1990. The ability to attract new jobs depends on concentration and not diversity, presumably because it takes some critical mass of workers in a given industry to generate the positive externalities that drive agglomeration.

As a robustness check I ran the regressions in Table 2.3 after removing the ten largest and ten smallest 3-digit HHI, MSA population, and percentage growth rates, respectively. My baseline specification is the logarithm of HHI, chosen for ease of interpretation, but I also ran the regression on the unlogged HHI. The coefficient concentration remained positive and significant, as shown in Table 2.6, below.

2.6 Conclusion

This paper examined the growth of private sector jobs across hundreds of cities in recent decades. I showed that economic concentration contributed to faster growth, a result that accords with much of the literature on agglomeration economies. However, I remind readers that reduced form regressions make policy implications ambiguous at best.

Table 2.1: Number of Private Sector Jobs Across the MSA Distribution

	1990		2018	
Rank	**MSA**	**Private Employment**	**MSA**	**Private Employment**
1	LA-Anaheim, CA MSA	4,786,862	NYC-Newark-, NY-NJ-PA MSA	6,399,146
100	Chattanooga, TN-GA MSA	97,422	Portland-South Portland, ME MSA	137,925
200	Lubbock, TX MSA	29,488	Amarillo, TX MSA	50,922
300	Yuba City, CA MSA	13,605	Great Falls, MT MSA	24,540

Table 2.2: Percentage Employment Growth by Decade

	1990-1999	**2000-2009**	**2010-2018**	**1990-2018**
Mean	29.90%	12.30%	18%	79.60%
Median	26.60%	7.70%	14.40%	55.60%
Max	141.40%	207%	170.50%	676.20%
	(St. George, UT MSA)	(Elizabethtown-Fort Knox, KY MSA)	(Salisbury, MD-DE MSA)	(Hinesville, GA MSA)
Min	-3.10%	-39.10%	-54.40%	-52.50%
	(LA-Anaheim, CA MSA)	(St. Joseph, MO-KS MSA)	(Lawton, OK MSA)	(Charleston, WV MSA)

Table 2.3: Effect of 3-digit NAICS Concentration on MSA Job Growth: 1990-2018

	Dependent variable:		
	growth_90_18		
	(1)	(2)	(3)
log(herf_3_90)	75.325***	48.848***	44.968***
	(9.695)	(11.094)	(11.599)
log(pop_90)		−15.788***	−15.734***
		(3.479)	(3.974)
man_loc_qot_90			−14.574**
			(6.315)
retail_loc_qot_90			32.751***
			(9.172)
hi_tech_loc_qot_90			7.028
			(9.028)
fire_loc_qot_90			1.296
			(9.135)
Constant	302.621***	394.545***	356.976***
	(29.008)	(34.747)	(49.558)
Observations	353	353	353
R^2	0.147	0.194	0.262
Adjusted R^2	0.144	0.190	0.249
Residual Std. Error	78.464 (df = 351)	76.362 (df = 350)	73.512 (df = 346)
F Statistic	60.363*** (df = 1; 351)	42.161*** (df = 2; 350)	20.442*** (df = 6; 346)
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 2.4: Effect of 3-digit NAICS Concentration on MSA Job Growth by Decade

	Dependent variable:		
	growth_90_99	growth_00_09	growth_10_18
	(1)	(2)	(3)
initial log(herf_3)	0.405	15.219***	19.505***
	(2.668)	(4.013)	(5.207)
Initial log(pop)	−4.085***	−7.978***	3.592***
	(0.914)	(1.313)	(1.365)
initial man_loc_qot	−2.806*	−4.278**	−0.277
	(1.453)	(1.781)	(1.517)
initial retail_loc_qot	3.178	1.109	12.086***
	(2.110)	(3.023)	(3.520)
initial hi_tech_loc_qot	2.267	1.450	−3.316
	(2.077)	(3.069)	(3.740)
initial fire_loc_qot	−3.876*	3.254	1.820
	(2.101)	(2.712)	(2.813)
Constant	76.337***	146.162***	22.145
	(11.400)	(17.346)	(20.180)
Observations	353	353	353
R^2	0.141	0.247	0.101
Adjusted R^2	0.126	0.233	0.085
Residual Std. Error (df = 346)	16.910	23.864	25.189
F Statistic (df = 6; 346)	9.468***	18.872***	6.448***
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 2.5: The Effect of Initial Diversity Across Time

	Dependent variable:			
	growth_90_18	growth_90_99	growth_00_09	growth_10_18
	(1)	(2)	(3)	(4)
log(initial_unique_NAICS)	-64.985^{***}	6.697^{*}	-27.955^{***}	-18.845^{***}
	(16.141)	(3.597)	(5.058)	(6.060)
log(initial_pop)	-3.146	-6.368^{***}	-0.243	4.683^{***}
	(5.939)	(1.323)	(1.834)	(1.765)
Constant	407.736^{***}	68.284^{***}	143.012^{***}	55.796^{***}
	(36.806)	(8.202)	(11.886)	(18.426)
Observations	353	353	353	353
R^2	0.187	0.112	0.259	0.027
Adjusted R^2	0.183	0.107	0.255	0.022
Residual Std. Error (df = 350)	76.692	17.090	23.531	26.046
F Statistic (df = 2; 350)	40.292^{***}	22.170^{***}	61.155^{***}	4.903^{***}

Note: *p<0.1; **p<0.05; ***p<0.01

Table 2.6: Robustness Checks

	Dependent variable:			
	growth_90_18			
	(HHI Outliers)	(Pop. Outliers)	(Growth Outliers)	(Unlogged HHI)
log(herf_3_90)	58.129***	28.245**	26.613***	557.869***
	(13.625)	(11.166)	(8.391)	(128.165)
log(pop_90)	−11.217***	−14.175***	−12.181***	−17.679***
	(4.013)	(4.174)	(2.844)	(3.742)
man_loc_qot_90	−12.208*	−16.169***	−15.770***	−15.827**
	(6.358)	(6.041)	(4.538)	(6.311)
retail_loc_qot_90	18.705**	21.810**	20.368***	32.148***
	(9.088)	(8.858)	(6.762)	(9.091)
hi_tech_loc_qot_90	1.685	4.094	14.683**	9.478
	(8.819)	(8.677)	(6.539)	(8.892)
fire_loc_qot_90	−4.594	0.516	−2.454	1.582
	(8.833)	(8.970)	(6.532)	(9.085)
Constant	367.940***	303.162***	273.235***	213.086***
	(50.665)	(49.928)	(35.370)	(46.118)
Observations	333	333	333	353
R^2	0.204	0.163	0.223	0.270
Adjusted R^2	0.189	0.147	0.209	0.257
Residual Std. Error	69.694 (df = 326)	67.718 (df = 326)	50.959 (df = 326)	73.116 (df = 346)
F Statistic	13.922*** (df = 6; 326)	10.553*** (df = 6; 326)	15.615*** (df = 6; 326)	21.289*** (df = 6; 346)
Note:				*p<0.1; **p<0.05; ***p<0.01

Figure 2.1: Economic Concentration and Total MSA Jobs: 2018

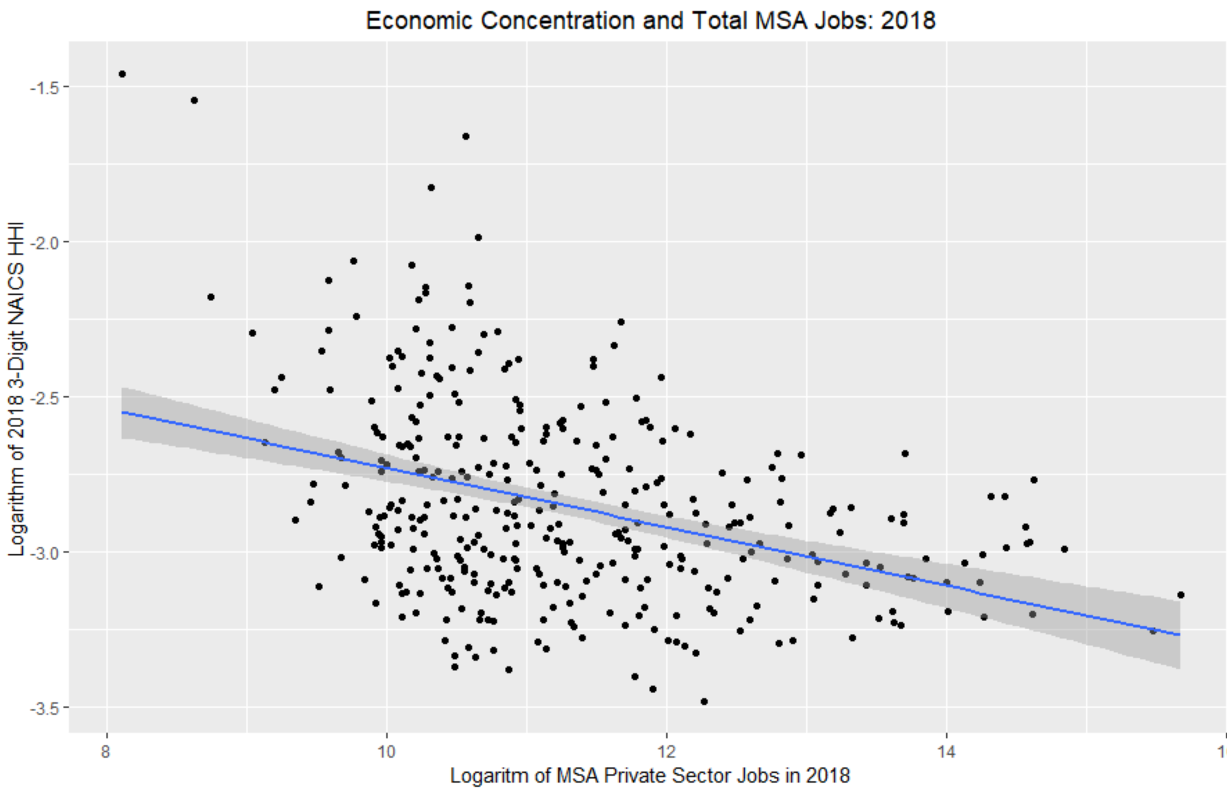


Figure 2.2: Industry Diversity and Total MSA Jobs: 2018

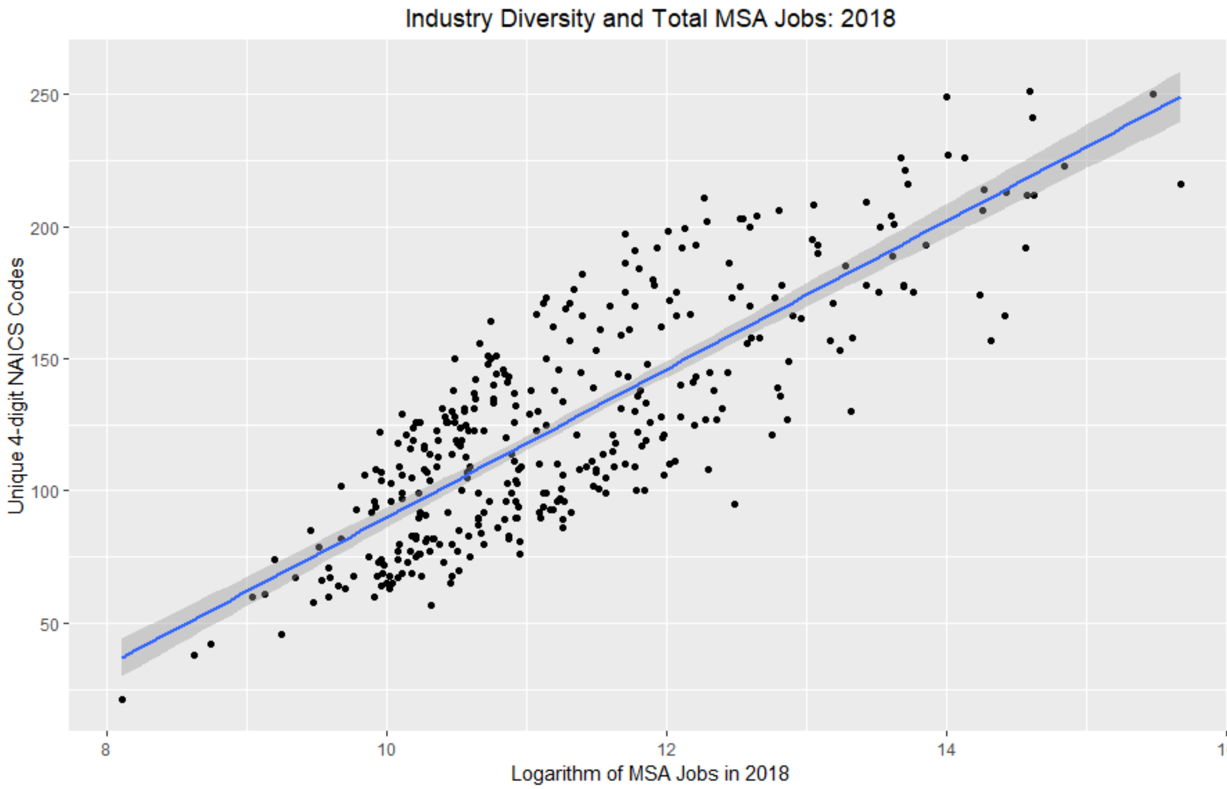
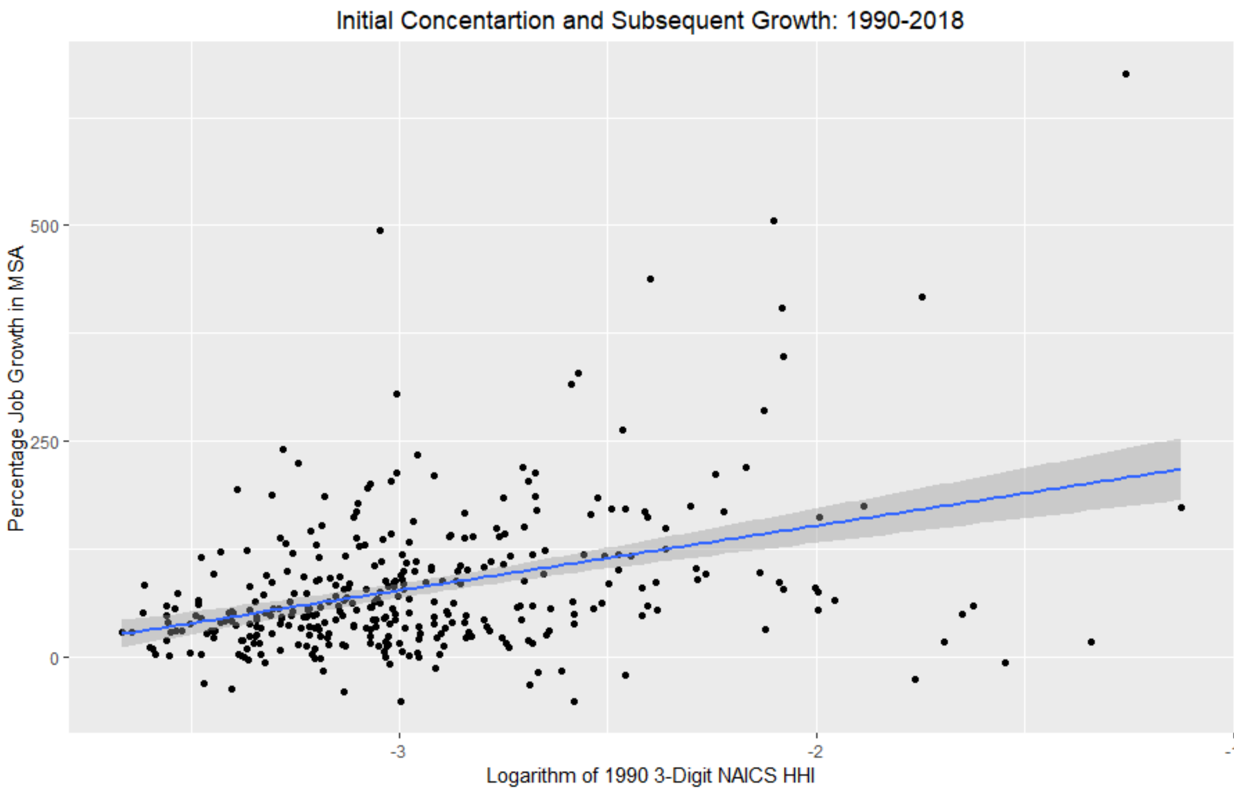


Figure 2.3: Initial Concentration and Subsequent Growth: 1990- 2018



CHAPTER 3

CONFLICT PROPENSITY AND ITS RELATIONSHIP TO CLIMATE CHANGE IN SUB-SAHARAN AFRICA (WITH ERIN BUNTING, ELIZABETH MACK, RICHARD A. MARCANTONIO, AMANDA ROSS AND ANDREW ZIMMER)

3.1 Introduction

The effects of climate change on temperature and rainfall may exacerbate existing political instability along fault-lines such as religion, ethnicity, and tribe, per Hsiang et al. (2013). Dell et al. (2012) found that higher temperatures reduce agricultural output, industrial output, and political stability. In warmer, drier climates, changes in rainfall may lead to more droughts. For example, Lesk et al. (2016) estimated that increased rates of drought and extreme heat result in a 7-11 percent cereal crop loss globally, causing food insecurity and other problems, especially in developing countries. Across the African continent, the physical and social effects of climate change are expected to be particularly acute given its dry climate and developing economics, per Buhaug et al. (2015) and Witmer et al. (2017).

Some argue that climate change is a driver of violent conflict, though this link has been debated in the literature and remains inconclusive, per Mach et al. (2019). Some argue that when water is not available people are more likely to engage in conflict, for example Burke et al. (2015), Homer-Dixon (1994), and Hsiang et al. (2013). Alternatively, others argue that an abundance of resources leads to more conflict, as people have more energy to engage in conflict and there are more resources to argue over. Given the climatological and other environmental trajectories for the African continent, it is important to understand these relationships and see how climate change may affect the likelihood of conflict.

In this paper, we examine how deviations from average rainfall affect the persistence of conflict in Sub-Saharan Africa. In general, areas at peace tend to remain at peace in the next period and areas in conflict are more likely to remain in conflict. However, areas do transition from conflict to peace and vice versa. Deviations in rainfall may affect the likelihood of transitioning from one state to another, specifically if one of the reasons for conflict relates to water. These findings complement those in Maystadt et al. (2014), who examined the effect of temperature anomalies on conflict and found that these deviations increase conflict.

We focus on Africa for several reasons. First, many countries in Africa are heavily dependent on agriculture. However, many African countries lack both well-developed irrigation systems and general infrastructure for household water access, per Salmon et al. (2015). As a result, Africa may be especially susceptible to changes in rainfall creating issues that lead to conflict. In addition, its position straddling the equator suggests that rising temperatures may have the largest effects in this region. Finally, institutions in many African countries tend to be lacking, per Michalopoulos and Papaioannou (2016). As a result, people may be more likely to resort to conflict, either violent or non-violent, to show their displeasure with the government.

To conduct our analysis, we first calculate the Markov transitional probabilities. Using data from the Armed Conflict Location and Event Database (ACLED), we consider different types of conflict across regions. Our focus is on deviations from the long-run trend in conflict because there is likely to be a steady state of conflict that people are used to and therefore we want to pick up deviations from this trend. Going forward, when we define peace and conflict we are referring to these deviations from the long-run average level of conflict. Using these transitional probabilities allows us to estimate the likelihood an area that was in peace (or in conflict) remains so, as well as the probability of transitioning between the two states. We consider these effects across the regions of Sub-Saharan Africa, as we believe regional differences will affect the likelihood of conflict in different areas.

We then use these transitional probabilities to estimate a logit model to look at how rainfall specifically affects the probability of transitioning from one state to another. Data on monthly rainfall was obtained from Climate Hazards Group Infrared Precipitation with Station (CHIRPS) data. We focus on the deviations from the long-run average in monthly rainfall versus aggregating rainfall to the annual or season level for two reasons. First because agriculture in a particular area is adapted to local conditions, making deviations more important than the absolute level. Second, focusing on monthly data allows us to see how the subtleties of climate change, such as a change in the length of the seasons, may change the likelihood of conflict in the area.

We find that areas at "peace" in one period, that is, below their long-run average level of violence, have a roughly 95 percent chance of remaining at peace next period. Areas above their long-run average violence level in one period revert to peace the next period roughly 64 percent of the time. What causes these shifts? After controlling for temperature and socioeconomic factors, we find that sufficient rainfall at the start of the dry season significantly lowers the odds that violence increases that year.

This paper contributes to the body of research exploring the relationship between environmental change – specifically climatic change and its impacts – and violent conflict. Anderson et al. (2017) linked bad harvests in Medieval Europe to the persecution of Jews. Jia (2014) showed that when farmers in 16th century China switched from wheat to the more drought-resistant sweet potatoes, their provinces saw fewer revolts. Bazzi and Blattman (2014) found that a positive shock to the price of a country's export shortened the duration of its conflicts. Kelley et al. (2015) demonstrate how a multi-year drought drove predominantly Sunni Muslim rural farmers into Shia Muslim controlled cities, exacerbating underlying social fractures that contributed to the still ongoing Syrian civil war. The changes to resources due to climate change have become more important as the effects of climate change become more rapid and acute, per IPCC (2018). Examining how changes in monthly rainfall affect the likelihood of transitioning into a state of conflict is important as uncertainty regarding future weather patterns continue to rise around the world.

Our research also complements a recent surge of research on Africa, specifically to the relationship between conflict and institutional quality. Nunn and Puga (2012) showed that in Africa rugged terrain increases income, the opposite of its effect on every other continent. They resolved this contradiction by showing that uneven land inhibited slave raiding, a practice whose long-run effects were even worse for subsequent growth than bad terrain. Brückner and Ciccone (2011) found that droughts in Africa undermined the credibility and capability of autocratic states and increased the odds of a transition to a more democratic form of government. Using the quasi-random borders imposed by Europeans that split over 200 ethnic groups, Michalopoulos and Papaioannou (2013) studied how regional factors play a decisive role in development and conflict. Along the same lines, Dickens (2018) showed that when rulers display ethnic favoritism, their patronage benefits geographic regions more than group members per se. Berman et al. (2017) measured the impact of mineral prices on conflict in Africa, finding that rising prices raise the risk of conflict by providing additional revenue for groups that controlled the mines, although the effect was not uniform across countries. Our paper adds to this research on Africa by showing that the in many places, the timing of abnormal weather patterns contributes to local conflict.

Finally, we contribute to a growing literature looking at the impact of climate change on society. Research on the wider social impact of climate change often falls into two categories. The first examines the economic impact of climate change, such as the Desmet and Rossi-Hansberg (2015) model of the global economic consequences of waring. More specific to Africa, Barrios et al. (2010) explored how changing rainfall patterns affected Africa's long-run growth. The second category explores the impact of climate change on migration, acting through general economic forces in Lilleør and Van den Broeck (2011), or conflict in particular in Reuveny (2007).

3.2 Conceptual Model

There are two different arguments regarding the relationship between changes in rainfall and conflict within a country. First is the scarcity argument, where insufficient resources drive conflict. Alternatively, there is the abundance argument, where the incentives of an excess amount of resources causes parties to fight over the resources. We discuss each of these theories and the evidence of both below.

Initially, many researchers argued that climate change and reductions in rainfall would lead to a higher incidence of conflict, per Homer-Dixon (1991). There are various reasons this negative relationship may occur. First, decreasing the supply of environmental resources, such as clean water and quality agricultural land, may lead to conflict over the allocation of these resources. Homer-Dixon (1994) showed this relationship through various case studies from the African continent. Almer et al. (2017) found that in unusually dry conditions, rioting increases, specifically in areas where there is more competition for water. Eriksen and Lind (2009) found a similar situation with regards to fewer water resources creating more conflict over water rights and access between farmers and herders. Couttenier and Soubeyran (2014) found a positive effect of drought on the likelihood of civil war. Another explanation is that agricultural production decreases when there is less rainfall and the reduction in food causes individuals to migrate to other areas in search of resources. This migration could lead to the outbreak of violence by exacerbating already existent social fracturing such as ethnic, religious, or other divides, per Kelley et al. (2015). Reuveny (2007) found that climate change induces migration and, in some cases, leads to conflict in the area where the migrants move. However, caution must be taken in seeing this link between migration and conflict as causal, per Abel et al. (2019).

Others argue that a negative relationship exists due to economic considerations. Specifically, if a decrease in rainfall reduces the profitability of an area, then the opportunity cost of conflict has decreased, per Collier and Hoeffler (1998). For example, if rainfall decreases and crops are ruined, farmers have less opportunity to profit through legal work. Therefore, the opportunity cost of conflict has decreased and more individuals may engage in conflict. Harari and Ferrara (2018) examined the effect of an agriculture-relevant weather shock on conflict. Their results suggest that a negative shock during the growing season contributes to more conflict in that area and the adjacent areas. Miguel et al. (2004) found a negative relationship between economic growth and civil conflict in 41 African countries. Alexandratos (2008) found that the rising price of staple crops led to protests and riots, specifically in urban areas. Roche et al. (2020) present a model which suggests there is a complex, and possibly nonmonotonic, relationship between climate variability and conflict.

Finally, extreme weather events could affect local government finances. Specifically, these events strain government revenues through a reduction in the tax base along with a simultaneous increase in the demand for services, per Benson and Clay (1998). Due to the reduction in finances, the government may not be able to keep its citizens happy and resist rebellions, per Fearon and Laitin (2003). Brückner and Ciccone (2011) found that negative rainfall shocks are followed by improvements in democratic institutions in Sub-Saharan Africa. Witmer et al. (2017) use a novel approach for predicting future violence across the African continent out to 2065 and argue that it is primarily the presence, density, and quality of institutional mechanisms and political rights for dealing with the strife caused by climate change that determines the likelihood of violent conflict arising. Overall, this research suggests that the deprivation of resources may lead to political changes or frustration towards that government that may result in conflict.

On the other hand, increases in rainfall may increase the likelihood of conflict. In fact, Theisen et al. (2012) noted that over the last ten to fifteen years there has been a drop in the frequency of civil wars in Africa, while temperatures and the incidence of droughts have simultaneously risen. Salehyan and Hendrix (2014) showed across a wide range of models that an abundance of water has a positive effect on the outbreak and sustainment of conflict, even after controlling for a range of characteristics such as the growth and level of GDP and population. Another argument for this positive relationship is that when rainfall decreases, individuals have less energy to devote to conflict, as people have other, more pressing issues just finding food, per Theisen et al. (2012).

From an economic standpoint, there are also counterarguments to the mechanisms discussed in the previous subsection. For example, if rainfall decreases and there are fewer crops, there are less resources to fight over, per Collier and Hoeffler (1998). Alternatively, a lot of rainfall may mean more resources, leading to increased conflict over these resources, per Klomp and Bulte (2013). Hendrix and Salehyan (2012) argue that there is a U-shaped pattern between conflict and rainfall, where too much rainfall is as disruptive as too little because the abundance of rain kills the crops just like the shortage. Specifically, they find that armed conflict is more likely to break out in wetter years, suggesting resource wars are more prevalent when there are resources to fight over. However, they also show that the other forms of conflict are strongly influenced by extreme positive and negative deviations from normal rainfall.

Alternatively, state capacity could create a positive relationship between rainfall and conflict. If rainfall increases resources, then the government may be better able to suppress a rebellion because they can keep the troops better fed and more engaged in supporting the government. Bazzi and Blattman (2014) examined the relationship between prices and conflict and found rising revenues improved the ability of governments to fight insurgents while also reducing the incentives of citizens to fight their government. This supports the conditional mobilization hypothesis which argues that rebellions are logistically demanding; to start and sustain conflict, armed groups must either receive support from (or extort) the local economy.

3.3 Data

We use the Armed Conflict Location and Event Dataset (ACLED) to measure conflict in Africa, introduced in Raleigh et al. (2010). ACLED geocodes the location of all reported conflict events in 55 countries in Africa and classifies each country into a region – Northern, Eastern, Western, Central, or Southern. Table 3.1 includes a list of all countries in ACLED as well as which countries are included in each region.

For our analysis, we draw upon information for each conflict situation on the event type, the exact date, and the exact location. The ACLED event types include protests, riots, remote violence/explosions, battles, strategic development, and violence against civilians. Further, the different types of events are broken into violent and nonviolent events, where nonviolent events are defined as protests and strategic developments. The other four event types are categorized as violent events. Table 3.2 provides the ACLED definitions for each type of conflict. Table 3.3 provides summary statistics. Figure 3.1 plots recorded events across Africa in 2018 for violent and nonviolent events.

We use the disaggregated, point level ACLED dataset on each individual event that occurred in Sub-Saharan Africa from 1997 to 2018. Rather than analyzing every single conflict data point, we utilize a raster-based methodology. First, an empty raster was developed at a 0.5° by 0.5° resolution. Then, based on an overlay of the raster with the conflict point data, every ACLED event was assigned to the appropriate grid cell. The assignment of points to grid cells was stratified in multiple fashions: (1) all conflict, (2) violent and nonviolent conflict events, and (3) by event type including battles, explosions remote violence, violence against civilians, riots, protests, and strategic development. The individual grid cell is our unit of analysis.

For each cell and conflict type, we calculated the average number of events across the 22 year time period. The long term average number of events was subtracted from the annual conflict to determine the yearly trend in the grid cell. Negative numbers indicate a year when ACLED events in a given cell fell below their long term mean and positive numbers indicate a year that events were above the long term mean. We normalize cells because some areas have more chronic violence than others. For example, cells in the middle of Cairo, cells in the middle of Harare, or cells straddling the border between Congo and Uganda tend to have higher levels of conflict. Examining these deviations from the long term mean allows us to better understand the changes in conflict relative to what is more normal for each area. For each cell and event type, we assigned a binary indicator for cells above (1) or below (0) their long-run average. A zero can be interpreted as "more peaceful than average" and a one as "more violent than average."

The Climate Hazards Group Infrared Precipitation with Station data (CHIRPS) is used to develop our annual, seasonal, and monthly precipitation measures. The CHIRPS data span 1981 to the near present at the 0.05 degree ($\sim$5km) resolution. Given the years in the ACLED database, we only use the CHIRPS data from 1997-2018. The rainfall data come from both geostationary infrared satellite data and in situ rain gauge observations which are then interpolated to produce precipitation raster data, per Funk et al. (2014). Figure 3.2, below, shows the mean and standard deviation for monthly rainfall across the CHIRPS Africa data set for 2018.

Both annual and monthly precipitation variables are used. For both temporal resolutions the total precipitation, precipitation variability, long term average annual precipitation, and annual deviation/anomalies are calculated. First, the monthly precipitation was calculated for each cell. Second, the per cell long-term average monthly precipitation was calculated. Next, the monthly precipitation was differenced from the long-term average monthly precipitation. Resulting variables above zero are indicative of higher than average rainfall and values less than zero are below average precipitation.

We focus on the monthly variation versus seasons in order to capture the transitional months between dry and wet seasons. For example, if climate change causes the dry season to last longer, that disruption in the transitional month may specifically be impactful on conflict in the region. Analysis at the monthly level allows us to examine temporal variation that would be lost using a seasonal or annual analysis.

We also include control variables for GDP, urban population from the World Bank (2020) and temperature from Matsuura and Willmott (2009).

3.4 Methodology

Markov transition matrices are used to measure the probability that a given area transitions into a different state or remains in its current state. In our context, we are examining whether areas transition into a higher or lower state of conflict than their long-run average. For each pair of years, such as 1997 and 1998, we find the total number of cells that were "more peaceful" in year one, then year two. If we observed 10 peaceful cells in 1997 and in 1998 nine of them were still peaceful and one transitioned to becoming more violent, the transition probabilities are (0.9, 0.1) where the first term indicates the odds that a peaceful cell stays peaceful and the second the odds it becomes violent. Similarly, if we observed 4 cells that were more violent than their long term average in 1997 and in 1998 one remained more violent and the other three became peaceful, we describe their transition probabilities as (0.75, 0.25) where the first term indicates the odds that a violent cell became more peaceful and the second the odds that it remained more violent. The Markov Transition Matrix for this example would be:

$$\begin{bmatrix} .9 & .1 \\ .75 & .25 \end{bmatrix}$$

Markov transition matrices are "memoryless" in that they only use information from the previous period. While this assumption is unrealistic for conflict, we address this by considering the deviation from the long term average for a given cell. For example, across the entire continent, a "peaceful" cell below its long-run average has a 95 percent chance of remaining peaceful the next year, and a 5 percent change to exceeding that average the next period.

We calculate these transition matrices for the level of total conflict across the African continent, as well as for each type of conflict because the persistence of conflict may vary by type of conflict. For example, we expect conflict such as explosions remote violence to be more random than protests. In addition, we break up our analysis by the different regions in Sub-Saharan Africa – East, Central, West, and Southern. This allows us to capture differences in culture and general trends that may be present.

Information from the Markov analysis is then incorporated into a panel logit model with the following specification:

$$Y_{i,t} = \alpha_i + \beta_1 STD\_DEV_{i,t} + \beta_2 MONDEV_{i,t} + \beta X_{i,t} + \epsilon_{i,t} \tag{3.1}$$

In this model, the dependent variable takes on value of 1 if in a given year a grid cell that was previously experiencing below average levels of conflict experienced a greater than average level of conflict as represented by the 22-year mean. If the grid cell remained at a below average level of conflict the next year, it received a value of 0. $STD\_DEV_{i,t}$ is the annual standard deviation of rainfall over 12 months in that cell-year. $MONDEV_{i,t}$ represents a vector of monthly dummy variables indicating whether rainfall was above the long-term mean in a particular month. In other words, a grid cell would receive a value of 1 if precipitation in that grid cell in a given month was above the long-term mean for that month. $X_{i,t}$ is the control variables included, specifically, GDP, urbanization, and temperature variables. α_i are grid-cell fixed effects and $\epsilon_{i,t}$ is a stochastic error term. Models were estimated with bootstrapped standard errors which have the advantage of not assuming a particular distribution to the error term.

Much of the early literature exploring the relationship between climate and conflict within a country started at the country level, such as Homer-Dixon (1994), Grey and Sadoff (2007), Siddiqi and Anadon (2011), Gleick (2014) or river basin level, such as Lowi (1993) and Zeitoun and Warner (2006). Ide (2015) considered 20 renewable resource disputes in the Global South from 1990 to 2010, but could not discern why some turned violent while other did not. However, Salehyan (2008) spelled out the problems with this line of research. First, violence does not occur over scarce resources per se, but rather as a result of failure to bargain over those resources. In other words, the level of analysis may create an endogeneity problem because Europeans imposed random borders on Africa. Michalopoulos and Papaioannou

(2016) showed that the former homelands for split groups tended to experience more violence. Overall access to water depends on the quality of government; a good government might spend more money on digging wells than on graft, for example. To solve that problem, Salehyan (2008) proposed we measure water stress through exogenous units of observation, arguing that small areas cannot drive national level policy changes. Therefore our analysis will focus on a more local level, looking at 0.5° by 0.5° grid cells.

This model was run for all the overall level of conflict, non-violent conflict, violent conflict, and each of the individual categories included in ACLED. In addition, we stratified our results across the regions in Sub-Saharan Africa – East, Central, West, and South. We do this because due to the geography of Africa straddling the equator, the wet and dry seasons vary notably across the continent. Aggregating all of Africa would likely to bias our estimates towards zero.

3.5 Results

Table 3.4 contains the transitional matrices produced from the Markov analysis. Across Africa, the likelihood that a grid cell at peace will remain in peace in the next time period is roughly 95 percent across all regions. So, the transition from a state of peace to conflict is rare, but Table 3.4 highlights that this transition probability does vary regionally. The probability of this occurring is 7.6 percent in East Africa and 5 percent in West Africa. These probabilities, while low, are higher than those for Central, Northern, and Southern Africa. Interestingly, the transition from conflict to peace is much more likely, and consistent across the five regions than is the transition from peace to conflict. So too is the likelihood that a grid cell in conflict will remain in conflict in the following year, much lower than is peace persistence from year to year.

If all conflicts are divided into violent and non-violent conflict, the Markov results illustrate subtle differences. First, peace persistence is higher in the analysis of non-violent conflict. The probability of a transition from peace to non-violent conflict is much lower than are the probabilities of for the transition from peace to violent conflict. Second, the probability of a transition from peace to violent conflict is higher in East Africa than the other four regions. The same is true for nonviolent conflict. Thus, for this time period, East Africa is more likely to transition to some type of conflict than are the other regions. Third, in this time period, violent conflict persistence is greater than is the persistence of non-violent conflict. Thus, grid cells in conflict are more likely to remain in some type of violent conflict than are grid cells where some type of non-violent conflict took place in the previous time period.

Table 3.5, below, presents the odds ratio for each independent variable for Sub-Saharan Africa as a whole and by region. An odds ratio is just that: a ratio of two probabilities. It communicates how much more or less likely the Y=1 outcome is after a 1 unit increase in the independent variable. For instance, if an annual temperature increase of 1 degree raises the baseline odds of violence from 1 percent to 1.2 percent, then the odds ratio would be .012/.01=1.2. So in Table 3.5, values above one indicate that an increase in the independent variable makes violence more likely, while a number below one indicates that an increase lowers the probability of violence. We determined the standard errors using the jackknife bootstrap.

The odds ratio for the standard deviation of rainfall is below one and highly significant. Why would a higher standard deviation make conflict less likely? Rainfall is right-skewed, so a drought during a cell's wet season would be the most likely reason the annual standard deviation fell. It would also make conflict more likely via the mechanisms discussed in Section 3.2. The large and significant odds ratio for the country's urbanization may be related to reporting: the international news outlets that form the basis of ACLED's data are based in cities. Riots and protests also occur disproportionately in urban areas.

The odds ratios more relevant for this paper are those associated with the monthly dummy variables indicating rainfall above the long run mean for that month in that cell. An odds ratio below one indicates that above average rainfall lowers the odds of conflict.

Moreover, we observe such odds ratios at the start of each region's wet season when farmers count on rain the most. In East Africa, the wet season runs February-May; the February odds ratio is 0.866. In West Africa, the wet season runs April-October; the April odds ratio is 0.974. In South Africa, the wet season runs October-April; the October odds ratio is 0.914, and the November odds ratio .791. We also observe odds ratios well above 1 for the November-March dry season in West Africa, indicating that seasonal anomalies matter for both wet and dry seasons. The Central Africa region has no wet or dry seasons, per se, although we do observe some months were deviations from the norm are associated with violence.

3.6 Conclusions

In this paper we considered the relationship between conflict and rainfall in Africa. We used Markov Transition matrices to show that since 1997, most places, most of the time, are relatively peaceful and likely to remain so. Unfortunately, an area that recently experienced a spike in violence could easily see that trend continue into the next period. We then focused on monthly rainfall to show that deviations from the historical norm increase the odds for violence, especially when rains fail to arrive early in the wet season.

These findings matter, not only for the countries studied, but for any area where farmers lack irrigation. The same government that failed to provide them with water infrastructure probably failed to delivery on other goods like education and healthcare. In that situation an income shock to farmers might well lead them to express their grievances, peacefully or otherwise.

Forecasts of global warming indicate that farmers could see seasonal abnormalities more and more in coming decades. If that happens, the conflict discussed above could spill across borders in the form of economic migrants and climate refugees.

Table 3.1: ACLED Regions

Central Africa:	Angola, Cameroon, Chad, Central African Republic, Democratic Republic of Congo, Equatorial Guinea, Gabon, Republic of Congo/Congo Brazzaville
East Africa:	Burundi, Djibouti, Eritrea, Ethiopia, Kenya, Madagascar, Malawi, Mozambique, Rwanda, Somalia, Tanzania, Uganda
North Africa:	Algeria, Egypt, Libya, Morocco, Sudan, Tunisia, Western Sahara
Southern Africa:	Botswana, eSwatini/Swaziland, Lesotho, Namibia, South Africa, Zambia, Zimbabwe
West Africa:	Benin, Burkina Faso, Gambia, Ghana, Guinea, Guinea-Bissau, Ivory Coast/ Cote d'Ivoire, Liberia, Mali, Mauritania, Niger, Nigeria, Senegal, Sierra Leone, Togo

Table 3.2: ACLED Conflict Types

Battles:	a violent interaction between two politically organized armed groups at a particular time and location
Explosions/Remote Violence:	one-sided violent events in which the tool for engaging in conflict creates asymmetry by taking away the ability of the target to respond
Violence against civilians:	violent events where an organized armed group deliberately inflicts violence upon unarmed non-combatants
Riots:	violent events where demonstrators or mobs engage in disruptive acts
Protests:	a public demonstration in which the participants do not engage in violence, though violence may be used against them
Strategic Developments:	contextually important information regarding the activities of violent groups that is not itself recorded as political violence, yet may trigger future events or contribute to political dynamics within and across states

Table 3.3: ACLED Events by Region: 1997-2018

Region	**Total**	**Battles**	**Explosions & Remote Violence**	**Violence Against Civilians**	**Riots**	**Protests**	**Strategic Deveoplemnts**
Central Africa	26,887	11,459	807	8,567	1,480	1,967	2,607
East Africa	56,790	20,443	6,560	15,259	4,444	6,424	3,660
North Africa	43,808	11,546	5,652	8,812	4,388	10,907	2,503
Southern Africa	17,514	280	83	4,934	5,156	6,618	443
West Africa	28,929	7,216	1,287	7,231	3,449	6,279	3,467

Table 3.4: Transition Probabilities

Region	Peace, Peace	Peace, Conflict	Conflict, Peace	Conflict, Conflict
		All		
Region	Peace, Peace	Peace, Conflict	Conflict, Peace	Conflict, Conflict
Central Africa	0.956	0.044	0.642	0.358
East Africa	0.924	0.076	0.616	0.384
North Africa	0.969	0.031	0.642	0.358
Southern Africa	0.957	0.043	0.673	0.327
West Africa	0.95	0.05	0.63	0.37
		Violent		
Region	Peace, Peace	Peace, Conflict	Conflict, Peace	Conflict, Conflict
Central Africa	0.959	0.041	0.641	0.359
East Africa	0.927	0.073	0.621	0.379
North Africa	0.971	0.029	0.604	0.396
Southern Africa	0.964	0.036	0.668	0.332
West Africa	0.953	0.047	0.646	0.354
		Non-Violent		
Region	Peace, Peace	Peace, Conflict	Conflict, Peace	Conflict, Conflict
Central Africa	0.982	0.018	0.702	0.298
East Africa	0.964	0.036	0.691	0.309
North Africa	0.983	0.017	0.732	0.268
Southern Africa	0.973	0.027	0.733	0.267
West Africa	0.976	0.024	0.633	0.367

Table 3.5: Panel Logit for all Types of Conflict

	Sub-Saharan	**East Africa**	**Central Africa**	**West Africa**	**South Africa**
Annual Std Dev of Precipitation	0.993***	0.993***	0.992***	0.997	1.001
	[0.000990]	[0.00132]	[0.00248]	[0.00221]	[0.00256]
Mean Annual Temperature	0.999***	1.001***	0.999***	0.999***	0.999
	[8.76e-05]	[0.000228]	[0.000175]	[0.000256]	[0.000771]
Variability in Annual Temp	1.000***	1	1.001***	1.001***	1.000**
	[7.16e-05]	[0.000180]	[0.000129]	[0.000277]	[0.000154]
National GDP	1.000***	1	1.000***	1.000*	1.000***
	[5.63e-11]	[3.81e-10]	[7.88e-10]	[1.02e-10]	[1.99e-10]
Percent Urban of Country	1.139***	1.193***	1.211***	1.173***	1.137***
	[0.00627]	[0.0117]	[0.0177]	[0.00952]	[0.0163]
January	1.110***	1.036	1.1	1.296***	0.845**
	[0.0286]	[0.0519]	[0.0660]	[0.0596]	[0.0635]
February	1.019	0.866***	0.998	1.156**	0.995
	[0.0242]	[0.0410]	[0.0446]	[0.0656]	[0.0663]
March	1.101***	1.051	1.285***	0.992	0.992

Continued on next page

Table 3.5 – *Continued from previous page*

	Sub-Saharan	**East Africa**	**Central Africa**	**West Africa**	**South Africa**
	[0.0281]	[0.0485]	[0.0695]	[0.0448]	[0.0643]
April	0.935***	1.005	0.916*	0.974	0.861**
	[0.0239]	[0.0420]	[0.0456]	[0.0531]	[0.0643]
May	0.981	1.027	1.062	0.975	0.741***
	[0.0252]	[0.0455]	[0.0555]	[0.0494]	[0.0652]
June	1.02	0.984	0.982	1.042	1.389***
	[0.0250]	[0.0400]	[0.0494]	[0.0566]	[0.0902]
July	1.031	0.993	1.077	1.097	0.964
	[0.0276]	[0.0492]	[0.0503]	[0.0673]	[0.0795]
August	1.050**	0.993	1.131**	0.894	1.172**
	[0.0234]	[0.0425]	[0.0576]	[0.0639]	[0.0855]
September	1.029	0.943	1.087*	0.915	1.190**
	[0.0247]	[0.0378]	[0.0527]	[0.0573]	[0.0901]
October	0.938**	0.952	0.885**	0.978	0.914
	[0.0238]	[0.0413]	[0.0461]	[0.0532]	[0.0662]
November	0.971	0.959	1.051	1.059	0.791***
	[0.0223]	[0.0391]	[0.0547]	[0.0567]	[0.0494]

Continued on next page

Table 3.5 – *Continued from previous page*

	Sub-Saharan	**East Africa**	**Central Africa**	**West Africa**	**South Africa**
December	1.153***	1.014	1.193***	1.227***	1.105
	[0.0267]	[0.0503]	[0.0620]	[0.0725]	[0.0766]
Observations	79,024	25,520	22,616	20,592	10,296

Figure 3.1: Distribution of Violent and Nonviolent Events across Africa in 2018

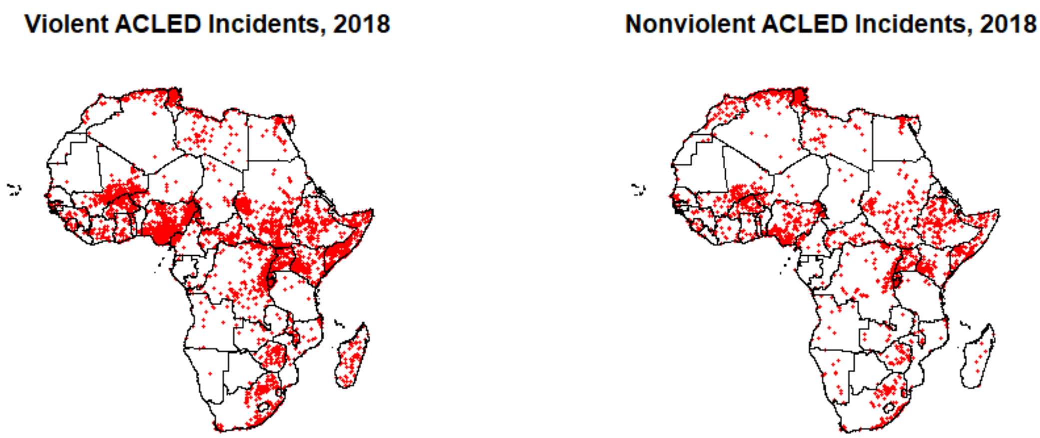
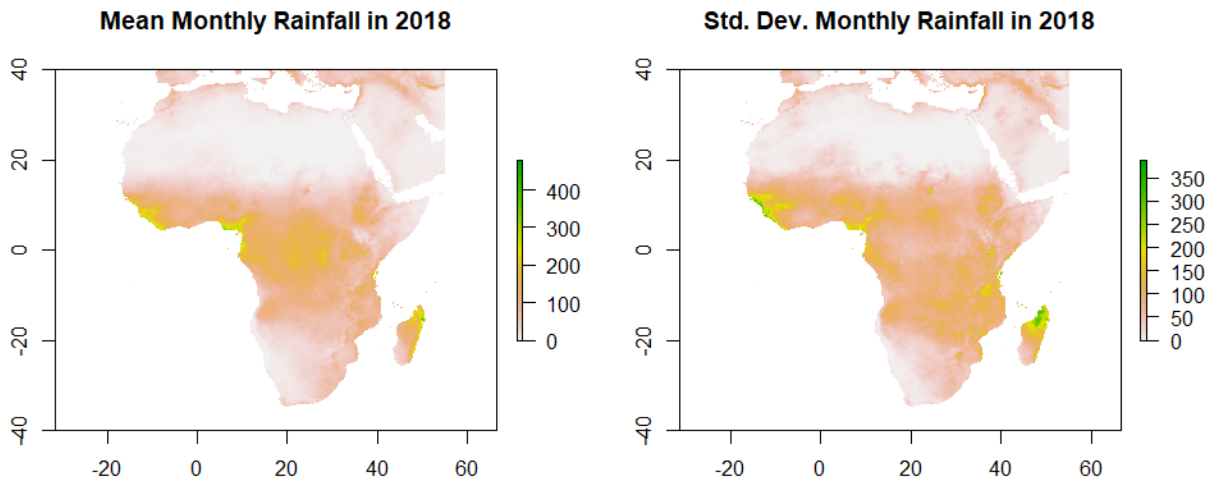


Figure 3.2: Mean and Standard Deviation of Rainfall in Africa in 2018



REFERENCES

Abel, G. J., Brottrager, M., Cuaresma, J. C., and Muttarak, R. (2019). Climate, conflict and forced migration. *Global Environmental Change*, 54:239–249.

Adelino, M., Ma, S., and Robinson, D. (2017). Firm age, investment opportunities, and job creation. *The Journal of Finance*, 72(3):999–1038.

Alexandratos, N. (2008). Food price surges: possible causes, past experience, and longer term relevance. *Population and Development Review*, 34(4):663–697.

Almer, C., Laurent-Lucchetti, J., and Oechslin, M. (2017). Water scarcity and rioting: Disaggregated evidence from sub-saharan africa. *Journal of Environmental Economics and Management*, 86:193–209.

Anderson, R. W., Johnson, N. D., and Koyama, M. (2017). Jewish persecutions and weather shocks: 1100–1800. *The Economic Journal*, 127(602):924–958.

Arellano, M. and Bover, O. (1995). Another look at the instrumental variable estimation of error-components models. *Journal of econometrics*, 68(1):29–51.

Barrios, S., Bertinelli, L., and Strobl, E. (2010). Trends in rainfall and economic growth in africa: A neglected cause of the african growth tragedy. *The Review of Economics and Statistics*, 92(2):350–366.

Bartik, T. J. (1991). *Who benefits from state and local economic development policies?* WE Upjohn Institute for Employment Research.

Bazzi, S. and Blattman, C. (2014). Economic shocks and conflict: Evidence from commodity prices. *American Economic Journal: Macroeconomics*, 6(4):1–38.

Benson, C. and Clay, E. (1998). *The impact of drought on sub-Saharan African economies: a preliminary examination*. The World Bank.

Berman, N., Couttenier, M., Rohner, D., and Thoenig, M. (2017). This mine is mine! how minerals fuel conflicts in africa. *American Economic Review*, 107(6):1564–1610.

Blank, J., Clary, C., and Nichiporuk, B. (2014). *Drivers of long-term insecurity and instability in Pakistan: Urbanization*. Rand Corporation.

Bloomberg 2020 (2020). All-in economy. Accessed 2 MARCH 2020, https://www.mikebloomberg.com/policies/all-in-economy.

United States Bureau of Labor Statistics (2020). Quarterly census of employment and wages handbook of methods. Accessed 2 MARCH 2020, https://www.bls.gov/opub/hom/cew/pdf/cew.pdf.

Brückner, M. and Ciccone, A. (2011). Rain and the democratic window of opportunity. *Econometrica*, 79(3):923–947.

Brunelle, C. (2013). The growing economic specialization of cities: Disentangling industrial and functional dimensions in the c anadian urban system, 1971–2006. *Growth and Change*, 44(3):443–473.

Buhaug, H., Benjaminsen, T. A., Sjaastad, E., and Theisen, O. M. (2015). Climate variability, food production shocks, and violent conflict in sub-saharan africa. *Environmental Research Letters*, 10(12):125015.

Burke, M., Hsiang, S. M., and Miguel, E. (2015). Climate and conflict. *Annu. Rev. Econ.*, 7(1):577–617.

Clarke, G., Li, Y., and Xu, L. C. (2016). Business environment, economic agglomeration and job creation around the world. *Applied Economics*, 48(33):3088–3103.

Collier, P. and Hoeffler, A. (1998). On economic causes of civil war. *Oxford economic papers*, 50(4):563–573.

Couttenier, M. and Soubeyran, R. (2014). Drought and civil war in sub-saharan africa. *The Economic Journal*, 124(575):201–244.

Dang, V. Q., Yang, Y., and Chan, K. S. (2018). Yen or yuan? the law of one price and economic integration in asia. *The World Economy*, 41(1):171–193.

Dell, M., Jones, B. F., and Olken, B. A. (2012). Temperature shocks and economic growth: Evidence from the last half century. *American Economic Journal: Macroeconomics*, 4(3):66–95.

Desmet, K. and Rossi-Hansberg, E. (2014). Spatial development. *American Economic Review*, 104(4):1211–43.

Desmet, K. and Rossi-Hansberg, E. (2015). On the spatial economic impact of global warming. *Journal of Urban Economics*, 88:16–37.

Diamond, R. (2016). The determinants and welfare implications of us workers' diverging location choices by skill: 1980-2000. *American Economic Review*, 106(3):479–524.

Dickens, A. (2018). Ethnolinguistic favoritism in african politics. *American Economic Journal: Applied Economics*, 10(3):370–402.

Du, K. (2017). Econometric convergence test and club clustering using stata. *The Stata Journal*, 17(4):882–900.

Duranton, G. (2007). Urban evolutions: The fast, the slow, and the still. *American Economic*

Review, 97(1):197–221.

Duranton, G. and Puga, D. (2004). Micro-foundations of urban agglomeration economies. In *Handbook of regional and urban economics*, volume 4, pages 2063–2117. Elsevier.

Duranton, G. and Puga, D. (2014). The growth of cities. In *Handbook of economic growth*, volume 2, pages 781–853. Elsevier.

Ellison, G., Glaeser, E. L., and Kerr, W. R. (2010). What causes industry agglomeration? evidence from coagglomeration patterns. *American Economic Review*, 100(3):1195–1213.

Engel, C. and Rogers, J. (1996). How wide is the border? *American Economic Review*, 86(5):1112–25.

Eriksen, S. and Lind, J. (2009). Adaptation as a political process: adjusting to drought and conflict in kenya's drylands. *Environmental management*, 43(5):817–835.

Fan, C. S. and Wei, X. (2006). The law of one price: evidence from the transitional economy of china. *The Review of Economics and Statistics*, 88(4):682–697.

Fearon, J. D. and Laitin, D. D. (2003). Ethnicity, insurgency, and civil war. *American political science review*, 97(1):75–90.

Funk, C. C., Peterson, P. J., Landsfeld, M. F., Pedreros, D. H., Verdin, J. P., Rowland, J. D., Romero, B. E., Husak, G. J., Michaelsen, J. C., Verdin, A. P., et al. (2014). A quasi-global precipitation time series for drought monitoring. *US Geological Survey data series*, 832(4):1–12.

Gabe, T. M. and Abel, J. R. (2012). Specialized knowledge and the geographic concentration of occupations. *Journal of Economic Geography*, 12(2):435–453.

Glaeser, E. L., Kallal, H. D., Scheinkman, J. A., and Shleifer, A. (1992). Growth in cities. *Journal of political economy*, 100(6):1126–1152.

Glaeser, E. L., Kerr, S. P., and Kerr, W. R. (2015). Entrepreneurship and urban growth: An empirical assessment with historical mines. *Review of Economics and Statistics*, 97(2):498–520.

Gleick, P. H. (2014). Water, drought, climate change, and conflict in syria. *Weather, Climate, and Society*, 6(3):331–340.

Grey, D. and Sadoff, C. W. (2007). Sink or swim? water security for growth and development. *Water policy*, 9(6):545–571.

Harari, M. and Ferrara, E. L. (2018). Conflict, climate, and cells: a disaggregated analysis. *Review of Economics and Statistics*, 100(4):594–608.

Hayakawa, K. (2009). First difference or forward orthogonal deviation-which transformation should be used in dynamic panel data models?: A simulation study. *Economics Bulletin*, 29(3):2008–2017.

Hellerstein, J. K., Kutzbach, M. J., and Neumark, D. (2015). Labor market networks and recovery from mass layoffs: Evidence from the great recession period. Technical report, National Bureau of Economic Research.

Hendrix, C. S. and Salehyan, I. (2012). Climate change, rainfall, and social conflict in africa. *Journal of peace research*, 49(1):35–50.

Homer-Dixon, T. F. (1991). On the threshold: environmental changes as causes of acute conflict. *International security*, 16(2):76–116.

Homer-Dixon, T. F. (1994). Environmental scarcities and violent conflict: evidence from cases. *International security*, 19(1):5–40.

Hsiang, S. M., Burke, M., and Miguel, E. (2013). Quantifying the influence of climate on human conflict. *Science*, 341(6151):1235367.

Ide, T. (2015). Why do conflicts over scarce renewable resources turn violent? a qualitative comparative analysis. *Global Environmental Change*, 33:61–70.

Intergovernmental Panel on Climate Change (2018). An ipcc special report on the impacts of global warming. Accessed 2 MARCH 2020, https://www.ipcc.ch/sr15/.

Jia, R. (2014). Weather shocks, sweet potatoes and peasant revolts in historical china. *The Economic Journal*, 124(575):92–118.

Kaplan, R. D. (2012). What's wrong with pakistan? *Foreign Policy*, 194:94.

Kelley, C. P., Mohtadi, S., Cane, M. A., Seager, R., and Kushnir, Y. (2015). Climate change in the fertile crescent and implications of the recent syrian drought. *Proceedings of the national Academy of Sciences*, 112(11):3241–3246.

Klomp, J. and Bulte, E. (2013). Climate change, weather shocks, and violent conflict: A critical look at the evidence. *Agricultural Economics*, 44(s1):63–78.

Lesk, C., Rowhani, P., and Ramankutty, N. (2016). Influence of extreme weather disasters on global crop production. *Nature*, 529(7584):84–87.

Lilleør, H. B. and Van den Broeck, K. (2011). Economic drivers of migration and climate change in ldcs. *Global environmental change*, 21:S70–S81.

Lowi, M. R. (1993). Bridging the divide: transboundary resource disputes and the case of west bank water. *International security*, 18(1):113–138.

Mach, K. J., Kraan, C. M., Adger, W. N., Buhaug, H., Burke, M., Fearon, J. D., Field, C. B., Hendrix, C. S., Maystadt, J.-F., O'Loughlin, J., et al. (2019). Climate as a risk factor for armed conflict. *Nature*, 571(7764):193–197.

Markusen, J. R. and Venables, A. J. (2013). *Functional specialization, sectoral specialization,* *and inter-city trade*. PhD thesis, The European Trade Study Group (ETSG) Working paper.

Matsuura, K. and Willmott, C. J. (2009). Terrestrial air temperature: 1900–2008 gridded monthly time series. *Center for Climatic Research, Dep. Of Geography, University of Delaware, Newark. http://climate. geog. udel. edu/ climate*.

Maystadt, J.-F., Calderone, M., and You, L. (2014). Local warming and violent conflict in north and south sudan. *Journal of Economic Geography*, 15(3):649–671.

Memon, R. (2005). Pakistan: internal migration and poverty reduction. *Migration Development and Poverty Reduction Asia*.

Michalopoulos, S. and Papaioannou, E. (2013). Pre-colonial ethnic institutions and contemporary african development. *Econometrica*, 81(1):113–152.

Michalopoulos, S. and Papaioannou, E. (2016). The long-run effects of the scramble for africa. *American Economic Review*, 106(7):1802–48.

Miguel, E., Satyanath, S., and Sergenti, E. (2004). Economic shocks and civil conflict: An instrumental variables approach. *Journal of political Economy*, 112(4):725–753.

Morshed, A. M., Ahn, S. K., and Lee, M. (2006). Price convergence among indian cities: A cointegration approach. *Journal of Asian Economics*, 17(6):1030–1043.

Moscarini, G. and Postel-Vinay, F. (2012). The contribution of large and small employers to job creation in times of high and low unemployment. *American Economic Review*, 102(6):2509–39.

Nunn, N. and Puga, D. (2012). Ruggedness: The blessing of bad geography in africa. *Review of Economics and Statistics*, 94(1):20–36.

Ottaviano, G. I. and Peri, G. (2006). The economic value of cultural diversity: evidence from us cities. *Journal of Economic geography*, 6(1):9–44.

Parsley, D. C. and Wei, S.-J. (1996). Convergence to the law of one price without trade barriers or currency fluctuations. *The Quarterly Journal of Economics*, 111(4):1211–1236.

Pakistan Bureau of Statistics (2017). Provisional summary results of 6th population and housing census-2017. Accessed 2 MARCH 2020, http://www.pbs.gov.pk/content/provisional-summary-results-6th-population-and-housing-census-2017-0.

Pakistan Bureau of Statistics (2018). Monthly price indices. Accessed 2 MARCH 2020, http://www.pbs.gov.pk/cpi-nb.

Pakistan Bureau of Statistics (2018). Methodology of price collection and computing price indices. Accessed 2 MARCH 2020, http://www.pbs.gov.pk/content/methodology-2.

Phillips, P. C. and Sul, D. (2007). Transition modeling and econometric convergence tests. *Econometrica*, 75(6):1771–1855.

Raleigh, C., Linke, A., Hegre, H., and Karlsen, J. (2010). Introducing acled: an armed conflict location and event dataset: special data feature. *Journal of peace research*, 47(5):651–660.

Reuveny, R. (2007). Climate change-induced migration and violent conflict. *Political geography*, 26(6):656–673.

Roche, K. R., Müller-Itten, M., Dralle, D. N., Bolster, D., and Müller, M. F. (2020). Climate change and the opportunity cost of conflict. *Proceedings of the National Academy of Sciences*.

Rosenthal, S. S. and Strange, W. C. (2004). Evidence on the nature and sources of agglomeration economies. In *Handbook of regional and urban economics*, volume 4, pages 2119–2171. Elsevier.

Rumi, R. A. (2016). Pakistan: Ungoverned spaces. *Sources of Tension in Afghanistan Pakistan: A Regional Perspective (2011-2016)*.

Salehyan, I. (2008). From climate change to conflict? no consensus yet. *Journal of Peace Research*, 45(3):315–326.

Salehyan, I. and Hendrix, C. S. (2014). Climate shocks and political violence. *Global Environmental Change*, 28:239–250.

Salman, A. (2015). Cane price, sugar surplus: Bitter solution to the sweet problem. *The Express Tribune*.

Salmon, J. M., Friedl, M. A., Frolking, S., Wisser, D., and Douglas, E. M. (2015). Global rain-fed, irrigated, and paddy croplands: A new high resolution map derived from remote sensing, crop inventories and climate data. *International Journal of Applied Earth Observation and Geoinformation*, 38:321–334.

Sanger, D. E. (2012). *Confront and conceal: Obama's secret wars and surprising use of American power*. Crown.

Sedláček, P. and Sterk, V. (2017). The growth potential of startups over the business cycle. *American Economic Review*, 107(10):3182–3210.

Siddiqi, A. and Anadon, L. D. (2011). The water–energy nexus in middle east and north africa. *Energy policy*, 39(8):4529–4540.

State Bank of Pakistan (2018). Inflation monitor. Accessed 2 MARCH 2020, http://www.sbp.org.pk/publications/Inflation_Monitor/index.htm.

Theisen, O. M., Holtermann, H., and Buhaug, H. (2012). Climate wars? assessing the claim that drought breeds conflict. *International Security*, 36(3):79–106.

Witmer, F. D., Linke, A. M., O'Loughlin, J., Gettelman, A., and Laing, A. (2017). Subnational violent conflict forecasts for sub-saharan africa, 2015–65, using climate-sensitive models. *Journal of Peace Research*, 54(2):175–192.

The World Bank (2016). *Pakistan development update : from stability to prosperity*. The World Bank Group.

The World Bank (2020). Indicators. Accessed 2 MARCH 2020, https://data.worldbank.org/indicator.

Zeitoun, M. and Warner, J. (2006). Hydro-hegemony–a framework for analysis of transboundary water conflicts. *Water policy*, 8(5):435–460.