Income and armed civil conflict: An instrumental variables approach

Christa N. Brunnschweiler* Päivi Lujala†

University of East Anglia Norwegian University of Science and Technology

July 2017

Abstract

The large empirical conflict literature has established that there is a strong negative link between economic variables and the onset of an armed civil conflict. However, it has been difficult to demonstrate a clear causality between poor economic performance and increased risk of conflict because of potential endogeneity issues, especially for large country samples. Most existing studies that analyse the causal links focus on the effects of economic growth on conflict, even though conventional conflict studies find the strongest relationship for income levels. In this article, we use three new exogenous instruments for income per capita, based on historical data for mailing times, telegram charges and urbanization rates. Using instrumental variables methods and global panel data for the period 1946-2014, we show that the negative effect of income per capita on the probability of conflict onset is consistently strong and larger than in conventional estimations using pooled ordinary least square regressions.

Keywords civil conflict, income, panel data, IV estimation

JEL codes D74, O10, C26

*Corresponding author: Christa Brunnschweiler, School of Economics, University of East Anglia, Norwich NR4 7TJ, United Kingdom. E-mail: c.brunnschweiler@uea.ac.uk.

†Päivi Lujala, Department of Geography, Norwegian University of Science and Technology (NTNU), 7491 Trondheim, Norway. E-mail: paivi.lujala@svt.ntnu.no.

1 Introduction

Economic variables have long been established in the empirical conflict literature as the most robust predictors of armed civil conflict (see e.g., Collier and Hoeffler, 2002, 2004; Fearon and Laitin, 2003; Hegre and Sambanis, 2006). However, although the link between economic variables and armed civil conflict seems strong, the causality is uncertain due to endogeneity issues, mainly arising from reverse causality and omitted variable bias. Several recent articles use 2SLS estimations, either with weather-based instruments (e.g., Bohlken and Sergenti 2010; Bergholt and Lujala 2012; Hodler and Raschky 2014; Sarsons 2015), or with commodity-price-based instruments (Brückner and Ciccone 2010; Bazzi and Blattman 2014; Berman et al. 2017), and analyze the effects of economic shocks on the likelihood of armed civil conflict, often with regional country-samples.[1]

This article contributes to this strand of literature by employing three new exogenous instruments for income per capita that can be used for a global sample of countries in armed conflict studies. Our instruments are based on historical data from the early 20th century on mailing times, telegram charges, and urbanization rates. They have several advantages over those used in other studies. First, they instrument income *levels*, while the vast majority of the existing papers that use two-stage least squares (2SLS) to explicitly model the intervening economic variable focus on income *growth*. The distinction is crucial since income *levels* have proven to be the more robust explanatory variable of the two in the conventional conflict literature. Second, our instruments are available for up to 179 countries, giving our results excellent external validity and allowing us to draw conclusions on global average causal effects. Finally, our instruments do not rely on climate-based or commodity-price data like most of the current instruments for economic conditions, but instead use new sources of exogenous variation.

To our knowledge, only two other contributions do not rely on either climate or commodity-price data. Djankov and Reynal-Querol (2010) use lagged savings rates as an instrument

[1] See the recent survey by Couttenier and Soubeyran (2015).

for income levels to find no robust effect of poverty on civil war; and Brunnschweiler and Lujala (2015) use the historical mailing times and telegram charges proposed in this article to instrument economic backwardness and income levels and explain episodes of – mainly nonviolent – mass social unrest. We focus instead on the link between income levels and armed civil conflict, and introduce one additional new instrument.

Using our three instruments in pooled 2SLS estimations, we confirm the importance of income levels for explaining the onset of armed civil conflict, and find that the magnitude of the effect is substantially larger than in conventional pooled ordinary least squares (OLS) estimations.

2 Methodology and data description

[Table 1 about here]

We follow the literature and explain the onset of armed *conflict* in year t in country i with income per capita levels and a vector of other covariates X according to:

$$conflict_{it} = a + \alpha_1 \cdot incomepc_{it} + \alpha_2 \cdot X_{it} + \epsilon_{it}, \tag{1}$$

where a is the constant term and ϵ the error term. We favor linear pooled OLS estimations, but also show our parsimonious baseline specification using the pooled probit model commonly used in the empirical conflict literature to take the binary-response nature of conflict data into account. Conflict is a zero-one dummy taken from the UCDP/PRIO Armed Conflict Database v.4-2014a (Gleditsch et al. 2002; Pettersson and Wallensteen 2015). It indicates the onset of a new armed civil conflict with over 25 battle related deaths or reactivation of a conflict after more than two years since the last observed fighting. Information for (ln) income per capita and (ln) population size comes from the Maddison dataset (Bolt and van Zanden 2014).[2] Other covariates include a polity measure and its square from

[2] We also used only the Penn World Tables (PWT 8.0, Feenstra et al. 2013) for the economic and

the Polity IV dataset p4v2015 (Marshall et al, 2013);[3] ethnic polarization (Montalvo and Reynal-Querol, 2005); mountainousness (Fearon and Laitin, 2003); (ln) oil reserves (Cotet and Tsui, 2013); the time in years since a country's independence or since 1945; a dummy for the post-Cold War period; a dummy for the colonial status in 1903 (own coding); and a dummy variable for historical conflict between 1816-1910 (own coding based on the Correlates of War dataset). Note that we lag all time-varying covariates except the post Cold War dummy by one year.

To take into account the likely endogeneity of income per capita, we follow Miguel et al (2004) and choose linear, pooled two-stage least squares (2SLS) as our approach[4] for the following first-stage estimation:

$$incomepc_{it} = c + \beta_1 \cdot I_i + \beta_2 \cdot X_{it} + \varepsilon_{it}. \tag{2}$$

We have a total of three different exogenous instruments I, which allows us to achieve a strong first-stage identification and to test for overidentifying restrictions.[5] Our dataset covers the period 1946-2010 and includes up to 9,253 country-year observations and 179 countries. Summary statistics are provided in Table 1.

Exogenous instruments. Our first instrument is based on mailing times in 1903 from either London or Washington, D.C. – whichever is faster – to the rest of the world. The two cities were chosen as they were the capitals of the world's most powerful economy at the time and of the world's soon-to-be most powerful economy, respectively. We used data on mailing times and distances for regular correspondence from Post Office Department (1903) and Post Office (1903),[6] supplemented by own calculations for missing cases. The "mailing speeds"

population variables. The sample loses three countries and over 1000 observations, but the results are remarkably similar (not shown).

[3] We use the "polity2" variable and add ten, so that it ranges from 0 (strong autocracy) to 20 (strong democracy).

[4] We do not use country fixed effects. See Beck and Katz (2001) for a theoretical argument against fixed effects in this context. A recent example in the conflict literature that does not use country fixed effects is given by Esteban et al (2012).

[5] Detailed descriptions of the instruments will be made available by the authors in an online codebook.

[6] We are grateful to Jenny Lynch from the US Postal Service and to staff at The Royal Mail Archive for

are calculated as miles covered per "mailing day". We then took the natural logarithm to construct our final measure, *Mailingspeed.* Mailing times are positively linked to economic development: not only did it take longer for correspondence to reach the more remote parts of the world; but at equal distances, letters reached a more developed and better-connected country before its "backwater" counterpart.

[Figure 1 about here]

The maps in Figure 1 show the travel time for letters, the distance the post covered, and the final mailing speed for present-day countries. For each map, the countries are divided into six equally-sized groups. Most of Europe, North America and the Caribbean, North Africa, and some parts of the Middle East – along the Suez Canal route – could be reached within ten days from either London or Washington, D.C. (see the map at the top). Southern parts of Africa and South America, as well as south-east Asia and Australia, were furthest away from Washington or London (the map in the middle). However, some of these places were reached relatively fast (see the bottom map), for example, South Africa – from whence letters were then dispatched by train to other inland colonies in southern Africa – and Australia. Mailing speed was slowest for some land-locked countries, in particular in the Sahel region and Central Asia.

The second instrument is international *Telegram* charges from Britain in (ln) pence in 1903, from Post Office (1903) and again supplemented by own calculations for missing cases. Telegram pricing principles were similar across the globe: they depended on distance and the number of words in the message, and also included labor costs right down to final delivery. A stated charge would apply to the first ten words together, and then to each single additional word (Ross 1928, Downey 2002). The resulting pattern meant that it cost less to send a telegram from London to Australia than to the West African coast, for example, so that telegram charges should be negatively related to income levels.

Our third and final instrument is the *Urbanization* rate (for towns 20'000 or larger) in

their help.

1920, taken mainly from UN (1969) and supplemented by information from McEvedy and Jones (1978), the Statesman's Yearbook (1922), and Bairoch (1988). Historical urbanization is strongly linked to the level of economic development at the time. For example, Acemoglu et al. (2002) use historical urbanization rates as proxies for economic prosperity and development. We expect that higher urbanization rates in 1920 are related to higher income per capita in more recent decades. Urbanization from a century ago is unlikely to affect conflict after WWII directly. Also, post-war urbanization patterns particularly in the developing world differed significantly from those seen earlier in the century (Bairoch 1988). So while cities may often be the epicenters of social unrest today, urbanization rates no longer have the strong positive link with economic development that they used to have.

[Figure 2 about here]

Figure 2 provides maps for telegram charges (top) and urbanization rates (bottom). Both maps show clear regional variations, but with different patterns. In 1903, it was most costly to send telegrams to South America and western Africa, while the urbanization rates were relatively high in South America and low in Africa.

We do not have the space to put forward all our theoretical arguments, but we believe that the three instruments satisfy the exclusion restriction. We think that our instruments offer an appealing alternative to those commonly used in the literature, based either on climate data or commodity prices. They are available for a large sample of countries across the world, ensuring excellent external validity of the results. The main downside is that our data (so far) exist only for one year at the aggregate country level. This limits the econometric approaches to those without country fixed effects.

[Table 2 about here]

In the reduced-form estimations in Table 2, we use variables from a parsimonious baseline specification including (ln) population and the dummy for a conflict in the previous year. Each instrument has the right sign and is highly significant when introduced on its own (columns 1-3). The instruments are not all individually significant when added together (column 4), but they are jointly highly significant (the null of joint insignificance is rejected at the 1% level).[7] Population and lagged conflict have the expected signs, though only population is significant. This first check supports the validity of our instruments. We also provide results of several statistical tests of instrument strength with the 2SLS estimation results below, and show some sensitivity checks.

3 Results

[Table 3 about here]

Table 3 first shows results using pooled Probit and OLS to confirm the standard finding of a highly significant negative coefficient for income per capita (columns 1-3). The remaining columns show the results of 2SLS estimations, initially using only one instrument at a time (columns 4-6), and finally using all three instruments together (columns 7-8). Income per capita is negative and highly significant across all specifications. Note that the magnitude of the coefficient increases in the 2SLS estimations when compared to the OLS estimations. The increase in the magnitude of the per capita coefficients in 2SLS estimations suggests that the upward bias in OLS estimations due to reverse causality is much smaller than the probable downward bias caused by omitted variables. The other explanatory variables are in line with the existing literature on armed civil conflicts.

Instrument validity. Looking at the first-stage information in the lower section of Table 3, all instruments have the expected sign and are highly significant both on their own (columns 4-6) and together (columns 7-8). Test statistics further confirm instrument strength and validity; tests based on Kleibergen-Paap rk Wald F statistics (not shown) also reject that we have weak instruments.

[Figure 3 about here]

As an illustration of the explanatory power of our instruments, Figure 3 depicts scatter-

[7] Angrist and Pischke (2009) discuss the importance of joint over individual significance in reduced-form estimations when there are several exogenous instruments.

plots for the first-stage regression fits of income per capita on our exogenous instruments, based on specifications (4)-(6) in Table 3. The plots illustrate the good performance of our instruments, though plot c) shows that Singapore (top-right) is a clear outlier in urbanization rates. Plot d) accordingly shows the same regression specification without Singapore, giving qualitatively unaltered results.

[Table 4 about here]

As a further test of the validity of our exogenous instruments, in Table 4 we add them in pairs (columns 1-3), which does not alter results and confirms IV strength. In column (4), we add a dummy variable for colonial status in 1903 (at the time the telegram charges and mailing speeds were measured), and a dummy for the occurrence of an armed conflict in the present-day country in the period between 1816-1910. All three of our instruments could conceivably be correlated with colonial status: colonies might on average be more prone to conflict after WWII as well as more developed, which would weaken our exclusion restriction. Similarly, historical conflict could be correlated both with development level and more recent conflict probability. The results clearly show that our instruments remain strong; that the main result for income per capita is not affected; and that neither former colonies nor countries that saw conflict further back in time were more likely to have an armed conflict onset after WWII.

4 Conclusion

We introduce three new instruments for income per capita and use these to test the relationship between income levels and the likelihood of armed civil conflict. We confirm that lower income levels increase the risk of conflict onset, and show that the effect is larger than with OLS.

References

Acemoglu D., Johnson S. , Robinson J. (2002), Reversal of fortune: Geography and institutions in the making of the modern world income distribution, Quarterly Journal of Economics 117(4): 1231-1294.

Angrist J., Pischke J.-S. (2009), Mostly harmless econometrics: An empiricist's companion, Princeton (NJ): Princeton University Press.

Bairoch P. (1988), Cities and economic development: From the dawn of history to the present, Chicago: University of Chicago Press.

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

Beck N., Katz J.N. (2001), Throwing out the baby with the bath water: A comment on Green, Kim, and Yoon, International Organization 55 (2): 487-495.

Bergholt D., Lujala P. (2012), Climate-related natural disasters, economic growth, and armed civil conflict, Journal of Peace Research 49(1): 147–162.

Berman N, Couttenier M. (2015), External shocks, internal shots: The geography of civil conflicts, The Review of Economics and Statistics 97(4): 758–776.

Bohlken A., Sergenti T. (2010), Economic growth and ethnic violence: an empirical investigation of Hindu–Muslim riots in India, Journal of Peace Research 47 (5): 589–600.

Bolt J., van Zanden L. (2013), The First Update of the Maddison Project; Re-Estimating Growth Before 1820, Maddison Project Working Paper 4, Available at http://www.ggdc.net/maddison project/publications/wp4.pdf. (Accessed 19 September, 2013).

Brückner M., Ciccone A. (2010), International commodity prices, growth and the outbreak of civil war in sub-saharan Africa, Economic Journal 120: 519–534.

Brunnschweiler C. N, Lujala P. (2015), Economic Backwardness and Social Tension, UEA Applied and Financial Economics Working Papers Series n.72.

Collier P., Hoeffler A. (2002), On the incidence of civil war in Africa, Journal of Conflict Resolution 46: 13–28.

Collier P., Hoeffler A. (2004), Greed and grievance in civil war, Oxford Economic Papers 56:563-95.

Cotet A. M, Tsui K. (2013), Oil and Conflict: What Does the Cross Country Evidence Really Show? American Economic Journal: Macroeconomics 5(1): 49–80.

Couttenier M, Soubeyran R. (2015), A Survey of the Causes of Civil Conflicts: Natural Factors and Economic Conditions, Revue d'Economie Politique 125 (6): 787-810.

Djankov S., Reynal-Querol M. (2010), Poverty and civil war: Revisiting the evidence, The Review of Economics and Statistics 92(4): 1035-1041.

Downey G. J. (2002), Telegraph Messenger Boys: Labor, Communication and Technology, 1850-1950, London: Routledge.

Esteban J., Mayoral L., Ray D. (2012), Ethnicity and Conflict: An Empirical Study, American Economic Review 102(4): 1310-1342.

Fearon J. D., Laitin D. D. (2003), Ethnicity, insurgency, and civil war, American Political Science Review 97(1): 75–90.

Feenstra R. C., Inklaar R., Timmer M. P. (2013), The Next Generation of the Penn World Table, available for download at www.ggdc.net/pwt

Gleditsch N. P., Wallensteen P., Eriksson M., Sollenberg M., Strand H. (2002), Armed conflict 1946–2001: A new dataset, Journal of Peace Research 39 (5): 615–637.

Hegre H., Sambanis N. (2006), Sensitivity analysis of the empirical literature on civil war onset, Journal of Conflict Resolution 50(4): 508–535.

Hodler R., Raschky P. A. (2014). Economic shocks and civil conflict at the regional level. Economics Letters 124: 530–533.

Marshall M. G., Gurr T. R., Jaggers K. (2013), Polity IV Project: Dataset Users' Manual, Center for Systemic Peace. http://www.systemicpeace.org/inscr/p4manualv2012.pdf (Accessed 19 September, 2013).

McEvedy C., Jones R. (1978), Atlas of world population history, London: Allen Lane.

Miguel E., Satyanath S., Sergenti E. (2004), Economic Shocks and Civil Conflict: An Instrumental Variables Approach, Journal of Political Economy 112 (4): 725-53.

Montalvo J. G, Reynal-Querol M. (2005), Ethnic Polarization, Potential Conflict, and Civil Wars, American Economic Review 95 (3): 796-816.

Pettersson T., Wallensteen P. (2015), Armed Conflict, 1946-2014, Journal of Peace Research 52 (4): 536–550.

Post Office (1903), Post Office Guide n.187, London: Post Office (Guidance & Services).

Post Office Department (1903), United States Official Postal Guide, January 1903, Philadelphia: United States of America Post Office Department.

Ross N. E. (1928), How to write telegrams properly, Little Blue Book n. 459, Girard (USA): Haldeman-Julius Publications.

Sarsons H. (2015), Rainfall and conflict: A cautionary tale, Journal of Development Economics 115: 62-72.

Statesman's Yearbook (1922), Statistical and historical annual of the states of the world for the year 1922, London: Macmillan.

UN (1969), Growth of the world's urban and rural population, 1920-2000, Department of Economic and Social Affairs, Population Studies n.44, New York: United Nations.

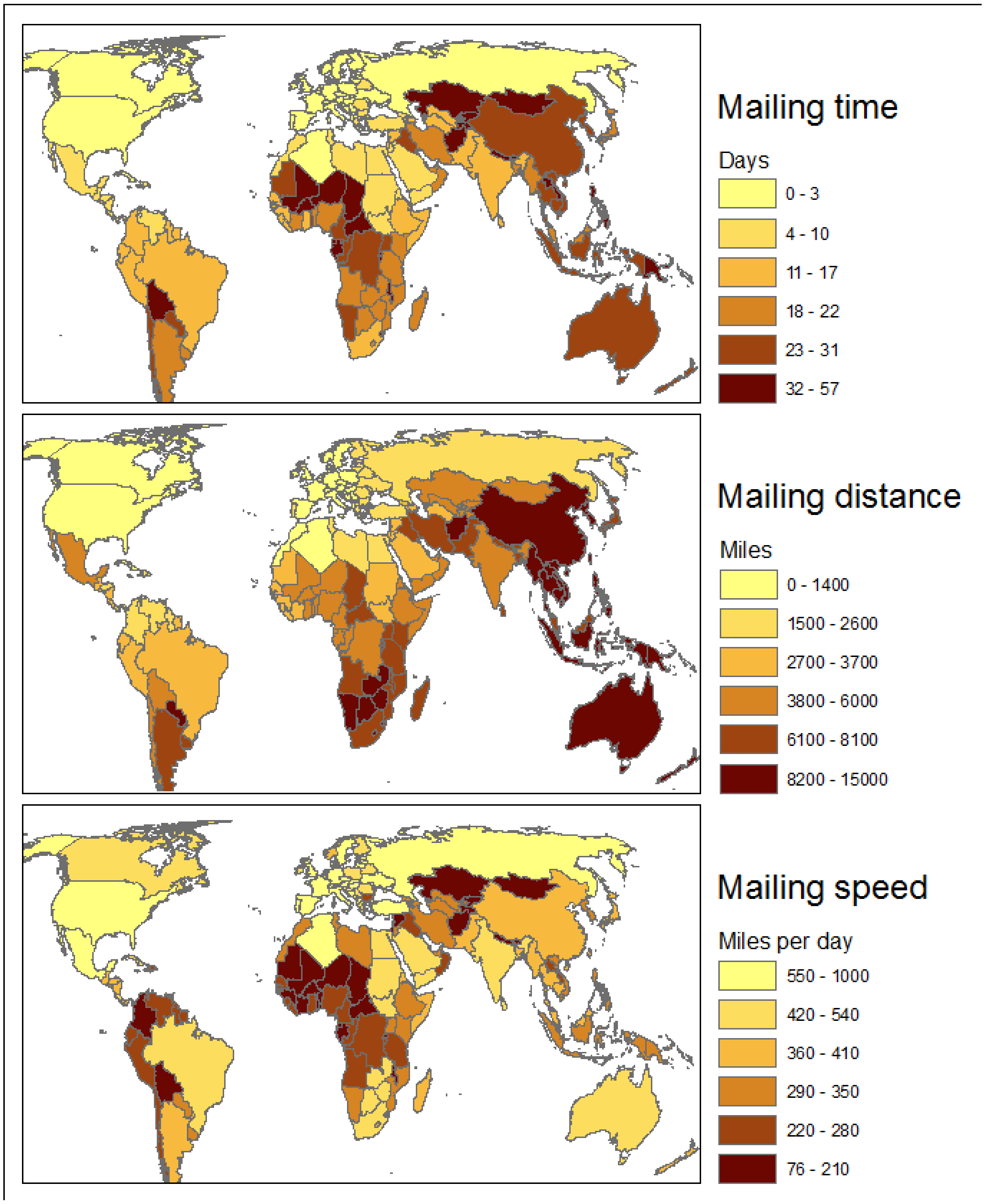

Figure 1: Mailing days, mailing distances, and final measure of mailing speed. Variables are divided into six equally-sized groups.

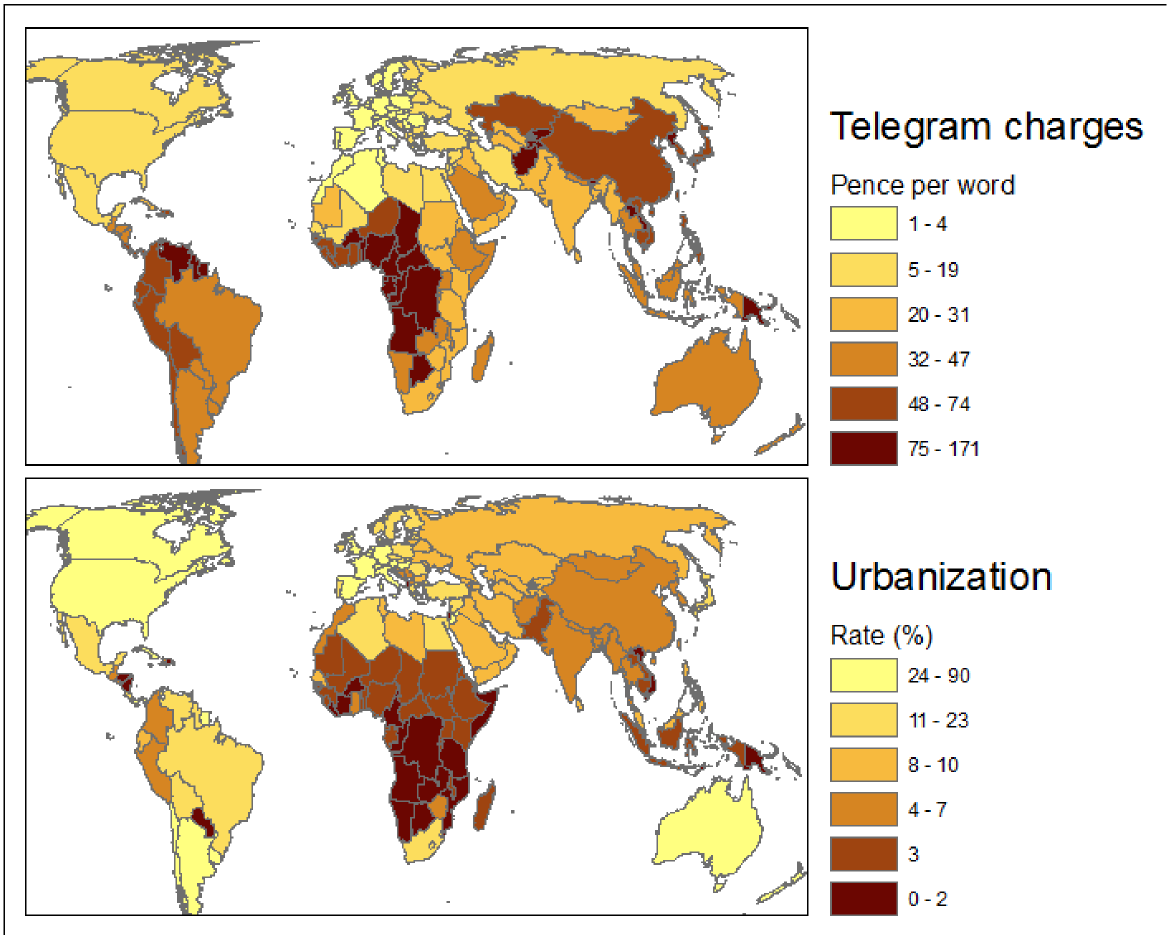

Figure 2: Telegram charges and urbanization rates. Variables are divided into six equally-sized groups.

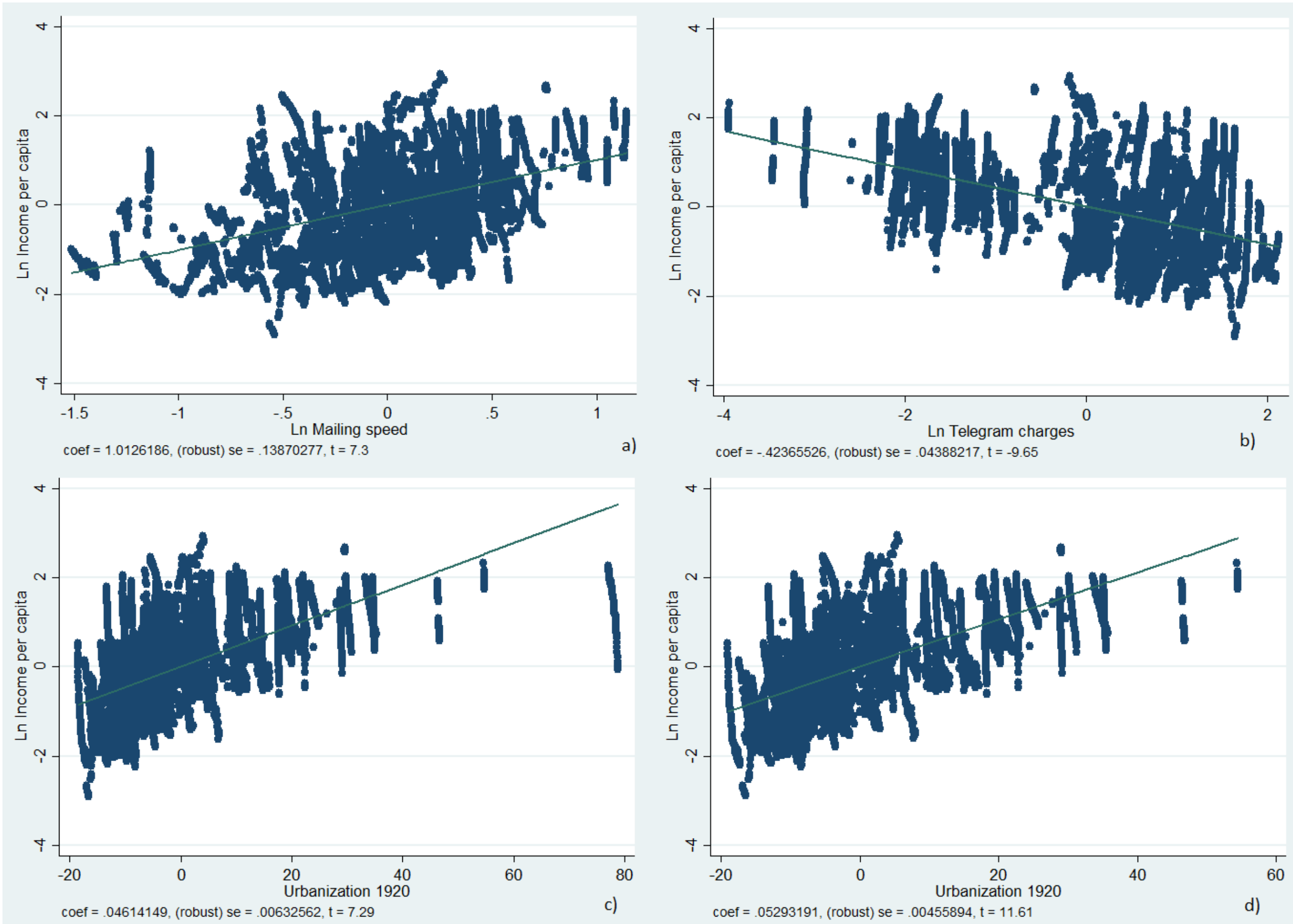

Figure 3: Scatter plots of first-stage results for exogenous instruments. Plots a)-c) correspond to specifications (4)-(6) in Table 2. Plot d) corresponds to specification (6) in Table 2 without Singapore. Lines show regression fits.

Table 1: Descriptive statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Conflict onset	9,253	0.034	0.181	0	1
ln Income p.c.	7,861	7.969	1.083	5.315	10.667
ln Population	8,545	9.02	1.57	4.824	14.107
Polity	9,155	10.608	7.460	0	20
ln Oil reserves	5,738	-4.579	5.274	-9.210	5.596
Mountainousness	8,888	18.288	21.665	0	94
Ethnic polarization	7,474	0.5191	0.242	0.017	0.982
Yrs since independence	9,253	29.683	18.63	1	69
Colony	9,253	0.397	0.489	0	1
Conflict 1816-1910	9,250	0.541	0.498	0	1
Telegram	9,253	38.137	33.373	0.5	171
Mailingspeed	9,253	348.819	155.060	75.702	1,000
Urbanization1920	9,253	12.690	14.065	0	90

Notes: For easier interpretation, *Telegram* and *Mailingspeed* are not reported in natural logs.

Table 2: Reduced-form estimations

	(1)	(2)	(3)	(4)
Mailingspeed	-0.0176***			-0.0056
	(-3.160)			(-0.872)
Telegram		0.0077***		0.0028
		(4.180)		(1.245)
Urbanization			-0.0009***	-0.0008***
			(-4.620)	(-3.686)
Conflict_{t-1}	0.0228	0.0216	0.0178	0.0171
	(1.490)	(1.446)	(1.222)	(1.163)
Population_{t-1}	0.0115***	0.0108***	0.0114***	0.0120***
	(3.652)	(3.457)	(3.670)	(3.850)
Observations	8,535	8,535	8,535	8,535
Countries	171	171	171	171
R^2	0.014	0.015	0.017	0.018

Notes: The dependent variable is onset of an armed civil conflict. All specifications include a constant term (not shown). S.e. are clustered at the country level. Robust t-statistics in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 3: Income and armed civil conflict onset

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Probit	OLS	OLS	2SLS	2SLS	2SLS	2SLS	2SLS
Income p.c.$_{t-1}$	-0.23***	-0.015***	-0.024***	-0.018***	-0.020***	-0.021***	-0.020***	-0.032***
	(-6.58)	(-6.02)	(-3.77)	(-3.39)	(-4.53)	(-6.14)	(-6.15)	(-3.17)
Population$_{t-1}$	0.14***	0.011***	0.0069	0.011***	0.011***	0.011***	0.011***	0.0055
	(4.40)	(3.26)	(1.46)	(3.34)	(3.33)	(3.35)	(3.35)	(1.33)
Conflict$_{t-1}$	0.12	0.014	-0.0056	0.013	0.011	0.010	0.011	-0.0095
	(0.96)	(1.00)	(-0.38)	(0.82)	(0.77)	(0.72)	(0.75)	(-0.67)
Polity$_{t-1}$			0.0043**					0.0034
			(2.01)					(1.45)
Polity$^2{}_{t-1}$			-0.0002					-0.0001
			(-1.50)					(-0.77)
Mountainousness			-0.0002					-0.00023
			(-1.35)					(-1.37)
Ethnic polarization			-0.0023					-0.0050
			(-0.16)					(-0.33)
Oil reserves$_{t-1}$			0.0033***					0.0041***
			(3.84)					(3.64)
Post Cold War			-0.012					-0.014*
			(-1.41)					(-1.68)
Yrs since independence			0.0005*					0.0007**
			(1.96)					(2.25)
Observations	7817	7817	4481	7817	7817	7817	7817	4481
Countries	161	161	94	161	161	161	161	94
R^2		0.019	0.025					
First-stage results								
Telegram					-9.65		-3.04	-1.71
Mailingspeed				7.30			2.96	3.91
Urbanization						7.29	6.18	8.58
F-statistic				53.3	93.2	53.2	43.2	30.7
Partial R^2				0.18	0.25	0.36	0.44	0.40
Hansen J p-value							0.75	0.72

Notes: The dependent variable is onset of an armed civil conflict. Column (1) shows pooled probit estimations; columns (2)-(3) pooled OLS; columns (4)-(8) pooled 2SLS. All specifications include a constant term (not shown). First stage information includes exogenous instruments' t-statistics, partial R-squareds, excluded instruments' F-statistics, and Hansen J statistic p-value. S.e. are clustered at the country level. Robust t-statistics in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 4: IV sensitivity analysis

	(1)	(2)	(3)	(4)
Income p.c.$_{t-1}$	-0.019***	-0.021***	-0.021***	-0.019***
	(-4.75)	(-6.06)	(-6.19)	(-4.32)
Population$_{t-1}$	0.011***	0.011***	0.011***	0.011***
	(3.34)	(3.35)	(3.35)	(2.96)
Conflict$_{t-1}$	0.012	0.011	0.011	0.010
	(0.79)	(0.74)	(0.75)	(0.74)
Colony				0.0049
				(0.50)
Conflict 1816-1910				0.0053
				(0.95)
Observations	7817	7817	7817	7817
Countries	161	161	161	161
First-stage results				
Telegram	-5.79	-4.56		-3.09
Mailingspeed	3.38		4.65	2.09
Urbanization		6.33	6.63	7.82
F-statistic	56.8	61.8	47.8	32.6
Partial R^2	0.29	0.42	0.41	0.33
Hansen J p-value	0.66	0.64	0.47	0.85

Notes: The dependent variable is onset of an armed civil conflict. All estimations are pooled 2SLS. All specifications include a constant term (not shown). First stage information includes exogenous instruments' t-statistics, partial R-squareds, excluded instruments' F-statistics, and Hansen J statistic p-value. S.e. are clustered at the country level. Robust t-statistics in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$