ARMA Modeling of Civil Violence: Case study of the Horn of Africa

Zhamilia Klycheva

Time Series Econometrics
Dr. Pierangelo De Pace
Claremont Graduate University
May, 2021

Abstract

Even though the number of civil conflicts has been decreasing over the past decades, the ongoing conflicts rage with new power. This is especially true for the countries in the Horn of Africa, particularly Somalia and Ethiopia. Time series analysis offers a unique analysis of the data changing over time and allows to explore the dynamics and model the unique events and their effect on future events. The current paper leverages simple ARMA models in the example of the two African countries, demonstrating that the parsimonious models alone can offer decent explanatory power and provide a good fit to the given data. Despite the similarity of the two states, each has its own unique history, resulting in peculiarities when it comes to the same modeling approach.

Background

Intrastate or Civil conflict is a whole branch of study under International Studies umbrella. Most of the literature from this field either focuses on studying the onset of the civil conflict or its severity (escalation) once it is initiated. In contrast to interstate conflict, civil conflict involves government and non-state actors. Civil war is an internal phenomenon that implies that even after the termination of war, the engaged parties will continue interacting as it might be hard to not to do so considering that actors reside in the same state.

Civil war studies go along with terrorism studies and are not often differentiated as both involve non-state actors and focus on the events occurring within a single state territory. The overall dynamics over time suggest that the number of civil conflicts has been seen a dramatic increase in 1990 suggested by Fearon and Laitin that it was due to the accumulation of protracted conflicts (Fearon and Laitin 2003). However the number has been recently decreasing since the beginning of a new century.

Unfortunately, despite the overall decrease in the number of civil conflict events, there are still countries with raging civil violence, where the Horn of Africa appears to be one of them. Somalia, a worn torn territory made it to the list of Failed states in 1991, however is now no longer considered a failed state, yet it has one of the highest Fragile State Index scores (Fragile State Index). Somalia has a rich history of violent events. Towards the beginning of the 21st century, Somalia experienced a series of events between Islamists and the Interim government backed by Ethiopia. The events unravel into a few years of civil conflict. Later on, the country falls under the hands of an terrorist organization Al-Shabaab that brings devastation and loss of civil lives till present day. The events involve the neighbor country Kenya that tries to fight the militant group. Currently, the country remains a fragile state, constantly involved in inter-clan clashes and attacks by Al-Shabaab.

Another country in the Horn of Africa currently experiencing civil conflict is

Ethiopia. Ethiopia shares a border with Somalia, Eritrea and Djibouti within the Horn states. It has also experienced a series of violent events, where early 2000s marked violent events with its neighbor Eritrea. This conflict was eventually resolved in 2018. During 2006, Ethiopia has interfered in Somalia's conflict and only withdrew few years later. Violent anti-government protests continue in the recent years, with the new government making drastic changes in 2018. Recent tensions involve the Tigray region based on ethnic tensions between ethnic Tigrayans that only comprise 6% of the population and the current government represented by a newly elected Oromo tribe prime-minister. These recent events picked up their pace and the number of violent events has seen its rise (BBC News).

This paper will focus on these two states of the Horn of Africa that have some overlapping history, yet unique events characteristic to each individual country.

Literature

There are multiple directions civil war literature takes that primarily involve panel data setup looking at multiple countries over time and predicting either the onset or severity of civil violence. These typical studies involve a number of political-economic variables such as economically motivated poverty, ethnic fractionalization, civil grievances, regime type, income per capita, education and so on (Collier and Hoeffler 2004, Fearon and Laitin 2003, Beson and Kugler 1998, Hegre et al. 2001). Multiple studies have conducted time series analysis on civil conflict. Brandt et al. (2010) conduct a forecasting analysis involving more recent literature on dynamic modeling of time series such as Markov-Switching Bayesian vector autoregression models (Brandt et al. 2010). Another technique used to predict ethnic conflict was used by Saideman et al. (2002) via Pooled Time Series emphasizing democratization and political institutions at predicting conflict (Saideman et al. 2002). Despite its relative simplicity, ARIMA modeling has also been used in this field of study. Lundgren (2020) studies the effect of conflict mediation via institutional intervention and the characteristics of such instutions via ARIMA modeling (Lundgren 2020). Schneider et al. (2012) have modeled one-sided civilian casualties to demonstrate them as an instrument of political leaders on their agenda and reciprocity by each side during Bosnian Civil War (Schneider et al. 2012). These models indicated strong results given a relatively parsimonious model, demonstrating a power of simple non-dynamic time series models.

Methodology

The current paper will use ARMA models to explain violent events in Somalia and Ethiopia to demonstrate the path dependence of the series. Based on this analysis, we will know whether ARMA models are an appropriate tool of working with this particular type of data and whether they can sufficiently explain the violent civil events per given state. Due to the data limitation, we did not obtain additional regressors as they are only available on an annual level which would substantially decrease the number of observations. Therefore, for the purposes of this paper, we offer a parsimonious version of this model as shown below:

$$SomViolence_t = c + \epsilon_t + \sum_{i=1}^{p} \phi_i SomViolence_{t-i} + \sum_{i=1}^{q} \theta_i \epsilon_{t-i}$$

$$EthViolence_t = c + \epsilon_t + \sum_{i=1}^{p} \gamma_i EthViolence_{t-i} + \sum_{i=1}^{q} \mu_i \epsilon_{t-i}$$

The ARMA models are estimated in STATA software via Maximum Likelihood Estimation. The choice of lags is shown in the Modeling section.

Data

As noted earlier in one of the definitions of the civil war, civilian casualties is a necessary condition. This paper will particularly model violent events that do not require a certain number of casualties, but rather fit into the "violent category". The data was provided by the Armed Conflict Location Event Data Project (ACLED) and the following violent categories were borrowed from the ACLED codebook.

Violent Events

Table 2: ACLED Event Types

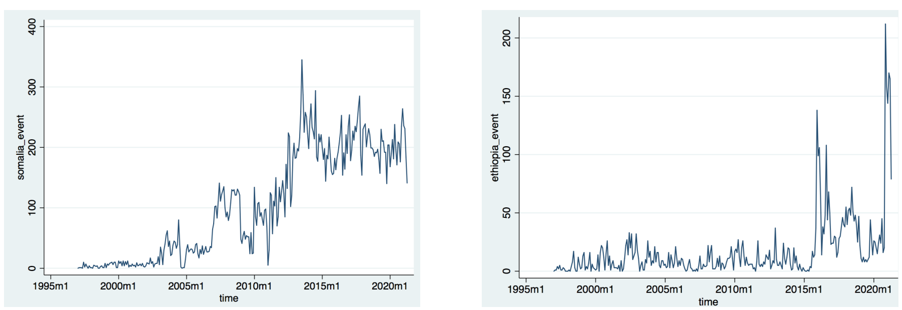
General	Event Type	Sub-Event Type
Violent events	Battles	*Armed clash*
		Government regains territory
		Non-state actor overtakes territory
	Explosions/Remote violence	*Chemical weapon*
		Air/drone strike
		Suicide bomb
		Shelling/artillery/missile attack
		Remote explosive/landmine/IED
		Grenade
	Violence against civilians	*Sexual violence*
		Attack
		Abduction/forced disappearance

Figure 1: Source: ACLED Codebook (2019)

The initial data came as a single event occurrence reported within a territory of a country. The data was then grouped on a monthly level. The time period is from 1997-2021, resulting in 292 observations. The data was limited to this time period as ACLED does not provide the data beyond 1997. As a note, the data was provided for academic use only.

Figures below show Time Series plots for the two countries of interest in the Horn of Africa, namely Somalia and Ethiopia. As we can clearly see, the data is not very smooth. This corresponds to all the events we discussed earlier happening within each state. Somalia has seen a relatively low level of events until 2006 as the state experienced transition in the government and no escalation occurred at this time. However during 2006-2009, Somalia deepened into a civil war, where Ethiopian troops fought on the behalf of the interim government against militias. When the situation stabilized, further in 2013 we observe another peak in the number of violent events triggered by the strengthening position of a terrorist group Al-Shabaab. While the number of events decreased since then, it rather plateaued and maintained a high level due to a strong presence of Al-Shabaab.

Ethiopia, the next door neighbor that shares commonalities in history has observed different violence dynamics. The country showed a relatively low and steady level of events all the way until 2016. In 2016, the country had a series of anti-government protests that visibly drove up the series. The events slightly slowed down shortly afetrwards, yet picked up the pace in 2018 when the new government came to power and the prime minister from Oromo ethnic group was elected. Recent spark in the events that is still ongoing till today is the conflict with the Tigray region in Ethiopia that has a historic significance, despite its small population percentage (Tigrayan ethnic group is only 6% of the entire population).



(a) Somalia

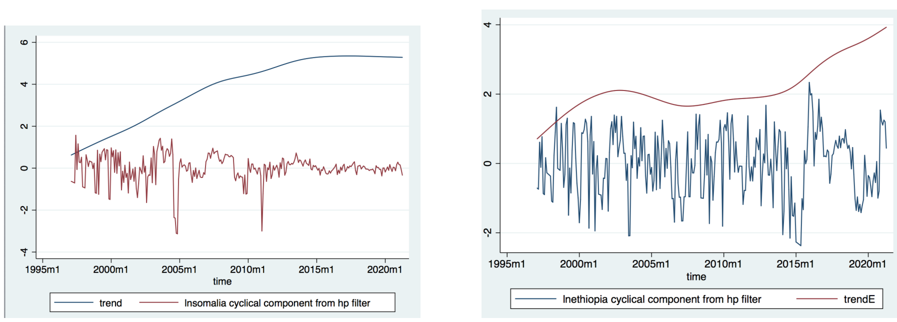
(b) Ethiopia

Figure 2: Conflict Time Series

Modeling and Results

In order to model the data using ARMA process, we need to ensure the variable is stationary and whether is passes few statistical tests. The very first step is taking a log of each of the series in order to smooth the variance. Second step is ensuring the data does not contain a unit root, and in the case of which applying a filter to remove a trend (if any). Lastly, once the data reaches an appropriate format, the series can be modeled and the evaluated based on information criteria and log likelihoods. Moreover, the best fitting model's residuals should be inspected to ensure they are stationary and can be relied upon.

To test for the presence of a unit root, we conduct an Augmented Dickey Fuller test. The results indicate that we fail to reject the Null hypothesis for both Somalia and Ethiopia Time Series, meaning a clear the presence of a unit root in both processes (See results in the Appendix).
Based on the results and the "trend" option in the Augmented Dickey Fuller test, both series contained trend. One of the common filters to detrend the data is a Hodrick-Prescott Filter (note that multiple filters have been tested that showed similar results, therefore we selected the Hodrick-Prescott as a widespread method).
The figures below show the detrended cyclical component series for Somalia and Ethiopia along with the extracted trends. As we can see, the extracted stationary series for Somalia (left) are still experiencing non constant variance over time. Ethiopian Time Series seem to be more of a constant nature relative to its neighboring country. We can note how different each series is, which will perhaps hint at different modeling techniques being more appropriate for one vs. the other.



(a) Somalia

(b) Ethiopia

Figure 3: Hodrick-Prescott Decomposition

Augmented Dickey Fuller test on the filtered time series above indicated no presence of a unit root, however the visual inspection of the ACF and PACF functions indicated some remaining structure for few periods, whereas the majority of the time the coefficients are not significantly different from zero (see Appendix). While this is not ideal, we will still go ahead and try modeling with these detrended series.

Somalia

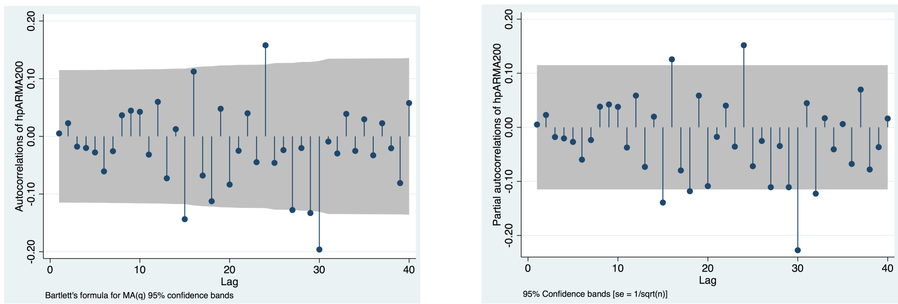
The figure below shows ARMA model outputs for the decomposed log Somalia conflict time series. We have tried running multiple models based on the initial

Autocorrelaiton and Partial Autocorrelation function plots for optimal numbers of autoregressive and moving average components. All the models in the table have significant beta parameters indicating the right fit of the AR and MA components. As we can see, ARMA(1,1) has a low value for Akaike Information Criterion(AIC) yet the highest value for the Bayesian Information Criterion (BIC). The middle column of an Autoregressive model of order 1 exhibits similar results with slighter higher values for AIC, yet lower value for BIC. However, it has the lowest Log Likelihood. Finally, column three or AR(2) shows the best results with the highest log likelihood, lowest AIC and second lowest BIC, indicating that this model is the most optimal for predicting the given data.

Variables	ARMA(1,1) hp_lnsomalia	ARMA(1,0) hp_lnsomalia	ARMA(2,0) hp_lnsomalia
L.AR	0.624***	0.479***	0.424***
	(0.0631)	(0.0316)	(0.0406)
L.MA	-0.187**		
	(0.0842)		
L2.AR Somalia			0.115***
			(0.0394)
Constant	-0.00441	-0.00303	-0.00446
	(0.0819)	(0.0723)	0.0823)
Sigma	0.559***	0.562***	0.558***
	(0.0137)	(0.0137)	(0.0136)
Observations	291	291	291
Log Likelihood	-243.72	-245.29	-243.35
AIC	495.45	496.58	494.697
BIC	510.14	507.60	509.39

Standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

As a next step, we need to inspect the residuals of the best model and see whether they are stationary. The plots below demonstrate ACF and PACF for ARMA(2,0) residuals which was selected as the best model earlier. Even though, the residuals are mostly not significant from zero, there are some periods that are significantly different from zero. Formal Q test showed that the Portmanteau (Q) statistic (standard test included with STATA) is equal to 0.0077, thus we fail to reject the Null verifying that the process is white noise.Though the results could have been better, they are not the worst in terms of passing the statistical test.



(a) ACF

(b) PACF

Figure 4: Diagnostic Plots for ARMA(2,0) Residuals

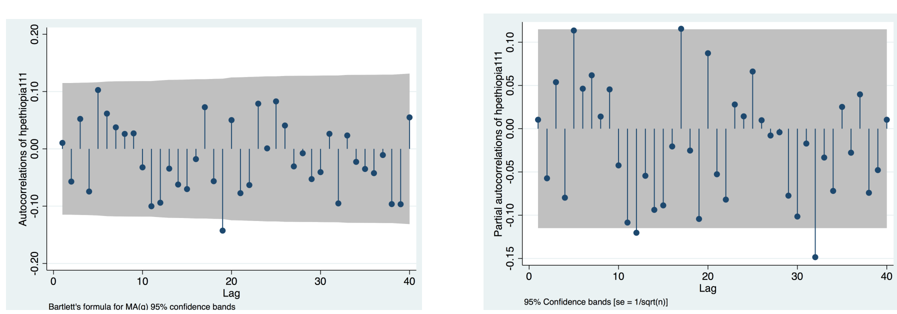
Ethiopia

Next, we are going to repeat the same process for Ethiopia Time Series. The table below shows the outputs for 3 ARMA models selected according to the similar principle as the table above. The maximum number of lags was limited to few periods only as the high number of lags showed no significance of parameters. Thus, the first model shows the highest Log Likelihood, lowest AIC, but highest BIC. Though the results might be tempting, the first autoregressive component is significantly not different from zero. Column 2 shows the lowest Log Likelihood value, yet lowest BIC. Finally, ARMA(1,1) demonstrates second highest Log Likelihood, slightly highest AIC and a slightly higher BIC. While there is no right or wrong answer, we would choose Model 3 as it had a higher log likelihood and almost equal AIC.

Variables	ARMA(2,2) hp_lnethiopia	ARMA (1,0) hp_lnethiopia	ARMA(1,1) hp_lnethiopia
L.AR	0.00428	0.497***	0.642***
	(0.222)	(0.0495)	(0.0891)
L2.AR Ethiopia	0.524***		
	(0.124)		
L.MA	0.461**	-0.197*	
	(0.228)	(0.111)	
L2.MA	-0.313***		
	(0.109)		
Constant	0.00140	-0.000817	-0.000411
	(0.120)	(0.0966)	(0.108)
Sigma	0.799***	0.809***	0.807***
	(0.0349)	(0.0358)	(0.0356)
Observations	291	291	291
Log Likelihood	-347.815	-351.53	-350.533
AIC	707.631	709.059	709.067
BIC	729.671	720.079	723.76

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

The residual autocorrelation and partial autocorrelation plots for ARMA(1,1) show that the residuals are not significant from zero for most of the time with a few small exceptions. A Portmanteau (Q) statistic of 0.0316 indicates that we fail to reject the Null confirming that the process is stationary. Clearly, this model is a better fit for Ethiopia relative to Somalia as one might have guessed by taking a look at the initial data. Indeed, while Somalia experienced much higher levels of volatility, Ethiopia had a moderately lower volatility which only increased towards the recent years.



(a) ACF

(b) PACF

Figure 5: Diagnostic Plots for ARMA(1,1) Residuals

Implications

As shown in the previous section, ARMA model demostrated similar results for both countries. Based on the evaluation of the residuals for each country model, it's clear that ARMA did a better job at predicting the violence for Ethiopia relative to Somalia (see ACF and PACF for both countries in Appendix). This goes back to the initial data peculiarities. While Somalia's best fitting model was ARMA(2,0), Ethiopia's best fitting model was ARMA(1,1). These choices can definitely be argued, however the logic was articulated earlier. It is important to note that given the simplicity, this model type is capable of fitting the data fairly well on the past data only. Yet, it was clear from the results that this type of data is best handled using other more sophisticated techniques as indicated in the previous literature that perhaps deal with the dynamic time series as the relationship between the series changes over time. Given the intricate nature of violence and as noted earlier, the multitude of factors contribute to civil violence and leveraging dynamic models would be interesting and worth exploring.

On the other hand, adding explanatory power into the current models by introducing additional regressors is another area that should be pursued. If the data limitation issue is resolved, adding new variables mentioned in the existing literature would add granularity and quality into even simple ARMAX model. Additional improvements into the modeling of variance and mean can be done via ARCH/GARCH processes as the data exhibits clearly heteroscedastic variances.
Finally, if the research question posed involves multivariate analysis, multivariate time series such as cointegration can be used to model more complex relationships between multiple time series. For the purposes of this study, one could have wondered whether there exists a long run equilibrium relationship between Somalian and Ethiopian violence time series. While these two countries share some common history, the relationship might not be robust enough to claim a potential link, although the exploration might be worthwhile in a new study.
As there is a limitless number of possibilities to explore, perhaps sticking to the most simple model could have a special value. This paper showed that modeling violence is indeed tricky and one should be extremely careful with the model selection, yet it's not impossible and definitely worth discovering.

Citations

ACLED. (2019). "Armed Conflict Location Event Data Project (ACLED) Codebook, 2019.

Benson M, Kugler J. Power Parity, Democracy, and the Severity of Internal Violence. Journal of Conflict Resolution. 1998;42(2):196-209. doi:10.1177/0022002798042002004

Brandt PT, Freeman JR, Schrodt PA. Real Time, Time Series Forecasting of Inter- and Intra-State Political Conflict. Conflict Management and Peace Science. 2011;28(1):41-64. doi:10.1177/0738894210388125

Carlsen, L., Bruggemann, R. Fragile State Index: Trends and Developments. A Partial Order Data Analysis. Soc Indic Res 133, 1–14 (2017). https://doi.org/10.1007/s11205-016-1353-y

Fearon, James D., and David D. Laitin. "Ethnicity, insurgency, and civil war." American political science review 97, no. 01 (2003): 75-90

Gerald Schneider, Margit Bussmann Constantin Ruhe (2012) The Dynamics of Mass Killings: Testing Time-Series Models of One-Sided Violence in the Bosnian Civil War, International Interactions, 38:4, 443-461, DOI: 10.1080/03050629.2012.697048

Håvard Hegre, et al. "Toward a Democratic Civil Peace? Democracy, Political Change, and Civil War, 1816-1992." The American Political Science Review, vol. 95, no. 1, 2001, pp. 33–48. JSTOR, www.jstor.org/stable/3117627. Accessed 1

Lundgren M. Causal mechanisms in civil war mediation: Evidence from Syria. European Journal of International Relations. 2020;26(1):209-235. doi:10.1177/1354066119856084

Ross, Will. "Tigray Crisis." BBC News, BBC, 18 May 2021, www.bbc.com/news/topics/cr2pnx1173dt/tigray-crisis.

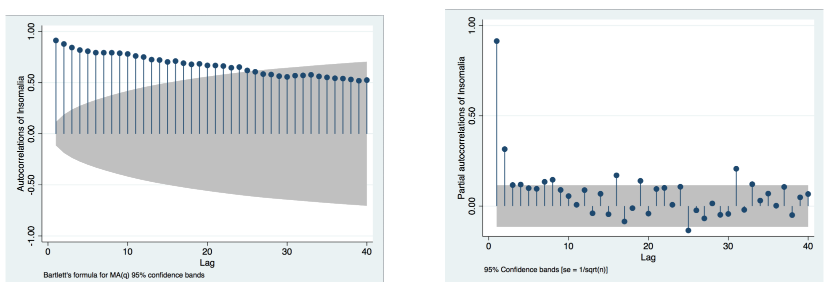
Saideman, Stephen M., et al. "Democratization, Political Institutions, and Ethnic Conflict: A Pooled Time-Series Analysis, 1985-1998." Comparative Political Studies, vol. 35, no. 1, February 2002, p. 103-129. HeinOnline, https://heinonline-org.ccl.idm.oclc.org/HOL/P?h=hein.journals/compls35i=99

Sambanis, Nicholas. "What Is Civil War?: Conceptual and Empirical Complexities of an Operational Definition." Journal of Conflict Resolution, vol. 48, no. 6, Dec. 2004, pp. 814–858, doi:10.1177/0022002704269355.

Singer, Joel David, and Melvin Small. Correlates of war project: International and civil war data, 1816-1992. Ann Arbor, MI: Inter-University Consortium for Political and Social Research, 1994.

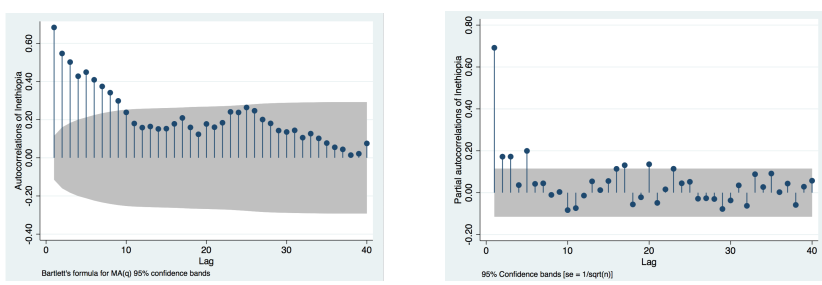
Appendix

The graphs below clearly indicate a presence of a unit root as the autocorrelation coefficients sustain over time and very slowly decay towards zero in the case of Somalia, but less apparent in the case of Ethiopia.



(a) ACF (b) PACF

Figure 6: Visually Inspecting Ln Somalia Time Series



(a) ACF (b) PACF

Figure 7: Visually Inspecting Ln Ethiopia Time Series

The test for checked using both "trend" and no "trend" option to check for the robustness of the results. In both cases we reject the Null, indicating the series contain unit root.

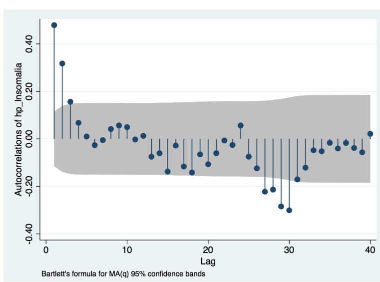
Augmented Dickey Fuller Test for Somalia Conflict

```
. dfuller somalia_event, lags(15)

Augmented Dickey-Fuller test for unit root         Number of obs   =       276

                               ---------- Interpolated Dickey-Fuller ---------
                  Test         1% Critical       5% Critical      10% Critical
               Statistic           Value             Value             Value
------------------------------------------------------------------------------
 Z(t)             -1.043            -3.458            -2.879            -2.570
------------------------------------------------------------------------------
MacKinnon approximate p-value for Z(t) = 0.7372
```

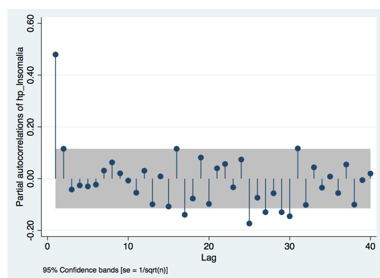
Augmented Dickey Fuller Test for Ethiopia Conflict

```
. dfuller ethiopia_event, lags(15)

Augmented Dickey-Fuller test for unit root         Number of obs   =       276

                               ---------- Interpolated Dickey-Fuller ---------
                  Test         1% Critical       5% Critical      10% Critical
               Statistic           Value             Value             Value
------------------------------------------------------------------------------
 Z(t)             -1.316            -3.458            -2.879            -2.570
------------------------------------------------------------------------------
MacKinnon approximate p-value for Z(t) = 0.6217
```

ACF and PACF plots for the detrended Time Series indicate that even though the coefficients seem to be not different from zero for the most part, there is a still some structure remaining, perhaps slightly less so for Ethiopia.

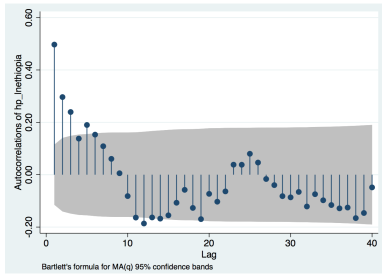


(a) ACF

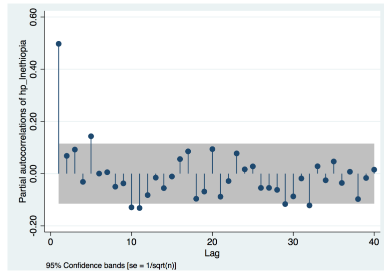


(b) PACF

Figure 8: Decomposed Time Series for Somalia



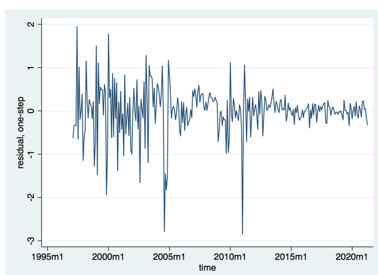
(a) ACF



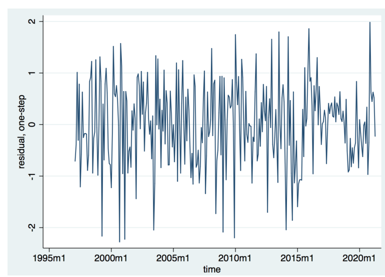
(b) PACF

Figure 9: Decomposed Time Series for Ethiopia

Residual Plots for each country's best model show that Somalian ARMA(2,0) suffer from some remaining structure, while Ethiopian ARMA(1,1) residuals exhibit much less volatility, yet still changing variance over time.



(a) Somalian ARMA(2,0)



(b) Ethiopian ARMA(1,1)

Figure 10: Residual Plots for the Best Fit Models