GDP MODELLING AND FORECASTING USING ARIMA: AN EMPIRICAL STUDY FROM INDIA

By

Varun Agrawal

Submitted to

Central European University

Department of Economics and Business

In partial fulfilment of the requirements for the degree of

Master of Arts in Economics

Supervisor: Professor Ariedo Muco

Budapest, Hungary

2018

Abstract

Gross domestic product (GDP) is one of most important macroeconomic indicators in an economy. This paper attempts model the time series of real GDP of Indian economy and subsequently develop forecast models to shed light on underlying data generation process. Using publicly available quarterly real GDP data from 1996, Quarter 2 to 2017 Quarter 2, I estimate various ARIMA models and calculate different forecasts. Results show that for the time period in contention, none of the ARIMA model proves to be strictly significant than other. I go ahead with AR(1) and MA(2) specifications and demonstrate the forecast evolution of real GDP and their residuals under different approaches. All the forecasts seem to be converging in the long run, though in the short run, big shocks like 2008 financial tend to cause a lot of divergence in the system.

JEL Classification: C53, E27

Keywords: GDP modelling, real GDP, Forecast, ARMA, Holt-winters

Acknowledgement

I am immensely grateful to professor Ariedo Muco for her continuous guidance and encouragement as my supervisor throughout the study. I would also like to extend my gratitude to professor Robert Lieli who helped me conceptualize and ideate the topic for this study and to Thomas Rooney for making sure the thesis version is readable. I'm also thankful all my teachers, the staff at Economics department, particularly Corinne Freiburger and department coordinator Katalin Szimler, and fellow students. Lastly, but not the least, I would like to express my appreciation to my family and friends for their unrelenting support and encouragement.

Table of Contents

List of figures

List of tables

"[O]nly by analysing numerous time series, each of restricted significance, can business cycles be made to reveal themselves definitely enough to permit close observation."

Burns and Mitchell, eds (1946), Measuring Business Cycles (p.11)

1. Introduction and Motivation

The purpose of this study is twofold. First, to ascertain which econometric model for time series analysis in ARIMA family fits the Indian GDP and Inflation quarterly data most accurately, and further do a forecast evolution exercise to validate the fit. Secondly, this study aims to undertake a comprehensive analysis of GDP forecasting and modelling for a developing country which could be used further as a template and reference for anybody interested in undertaking the ambitious exercise of modelling an emerging country's GDP.

Given the importance of GDP data in modern society, from becoming an election narrative to influencing the global commodity market, it is imperative that ample research should be done for GDP modelling for any country particularly for developing countries such as India. However, the reality seems far from it. As observed in Das and Do (2014), the number of research articles in top 5 economics journal[1] talking about developing countries has been around 2% , for example only 39 papers focused on India from 1985 to 2004 whereas 2,383 paper focused on United States.

Such dearth of research in GDP modelling for country like India is surprising and serves as my main motivation to undertake this study which is to come with a holistic approach towards modelling of Indian GDP numbers setting a template which could be used

[1] Top 5 lists here as AER, Econometrica, QJE, JPE and ReStud.

further as more data becomes available. My secondary contribution is to demonstrate the approach for modelling irrespective of the results so that more general interest could be generated in such topics. I start with looking at Indian GDP numbers and applying preliminary filter . Then, I proceed with few forecast methods to give an idea about the underlying process. Eventually, I use ARIMA family models to estimate the GDP series and the results does not seem conclusive and hence does not point to any single model that fits the data explicitly. Hence keeping in mind my other motivation, I go further with MA (2) and AR(1) specification and demonstrate the forecasting mechanics. I used quarterly GDP data from 1996 to 2017 made available by Indian Central Bank and National Sample survey organization, the government arm for statistical database collection across the country.

My biggest contribution to the literature is the claim that this is the first study analysing the quarterly data for Indian GDP series to estimate and forecast the data process. My results demonstrate that for the data in consideration, there is no conclusive proof that pins down the projection of Indian GDP growth series to an ARIMA model and hence forecasting methods are not too informative either. However, the methodological approach of undertaking this exercise makes me confident of the contribution I can make to the literature given mostly all previous studies focus on yearly data.

Following the introduction, review of literature[2] is presented. In next section I give an overview of the data sources used. In next section thereafter, ARIMA model and assumptions are detailed and then detailed modelling, analysis and results are presented. Finally, the conclusion section completes the main structure of the thesis followed by appendix and reference.

---

[2] Becketti (2013) gives a comprehensive overview of evolution of ARIMA modelling .

2. Review of Literature

Keeping in sync with the goal of this I study, I limit the review of literature to empirical ideas and research done for modelling GDP series of developing countries, particularly India. I do not go into any details about the theoretical constructs regarding GDP and components, forecasting approaches literature but mainly focus on empirical reviews.

The use of ARIMA models for GDP forecasting was started with the seminal Box and Jenkins (1976) paper. In Indian context, the first study undertaking GDP modelling using ARIMA is Maity and Chatterjee (2012) where they find that only in one period across the GDP series (1951-2011) ARMA terms were significant. They fit a simple ARIMA (1,2,3) . Their results on forecasting, by their own admission, suggest an upward trend but growth rates showing opposite trend for future periods. Changle and Matharu (n.d.) conduct a similar study with same dataset and found that forecasts showed positive trends. For a developed country prospective Zhang Haonan (2013) observe using Sweden data for 16 years that 1st order ARMA showed the most significant results. For a better suited comparison to Indian context , Zakia(2014) used Pakistan GDP data and found ARIMA (1,1,0) to be the best fit using quarterly numbers which comes closest to the approach demonstrated in this study. Dritsaki (2014) used Greece GDP data from 1980 to 2013 to fit an ARIMA (1,1,1) model forecasting values for three years in future. She found that the forecast to be showing an upward trend in GDP growth numbers.

The most comprehensive study done estimating the GDP series using ARIMA and forecasting is Waboma et al (2015) which looks at data from Kenya and find out that ARIMA (2,2,2) fits their data best with forecast in sample being 5% close the actual numbers.

3. Data description

I use four types of data variable for this study. One is quarterly date[3] where 1996:2 would mean second quarter of year 1996. Quarter 1, until specified covers the date range from 1st January to 31st March of respective year. I used nominal GDP, real GDP based on year 2000 prices and a measure of GDP deflator. The reason for choosing base year as 2000 was reliably available GDP series . GDP numbers are in billion rupees[4]. The original source of the data is Reserve bank of India and Central Statistical Office, New Delhi which collects data through primary surveys. In this study, I use a dataset[5] complied by FRED,St. Lousi, USA built on original dataset available on Indian Central bank website. Formula for calculating GDP used by the data source is rather a simple one:

GDP (Factor Cost) = GDP (Market Price.) -Indirect Taxes + Subsidies.

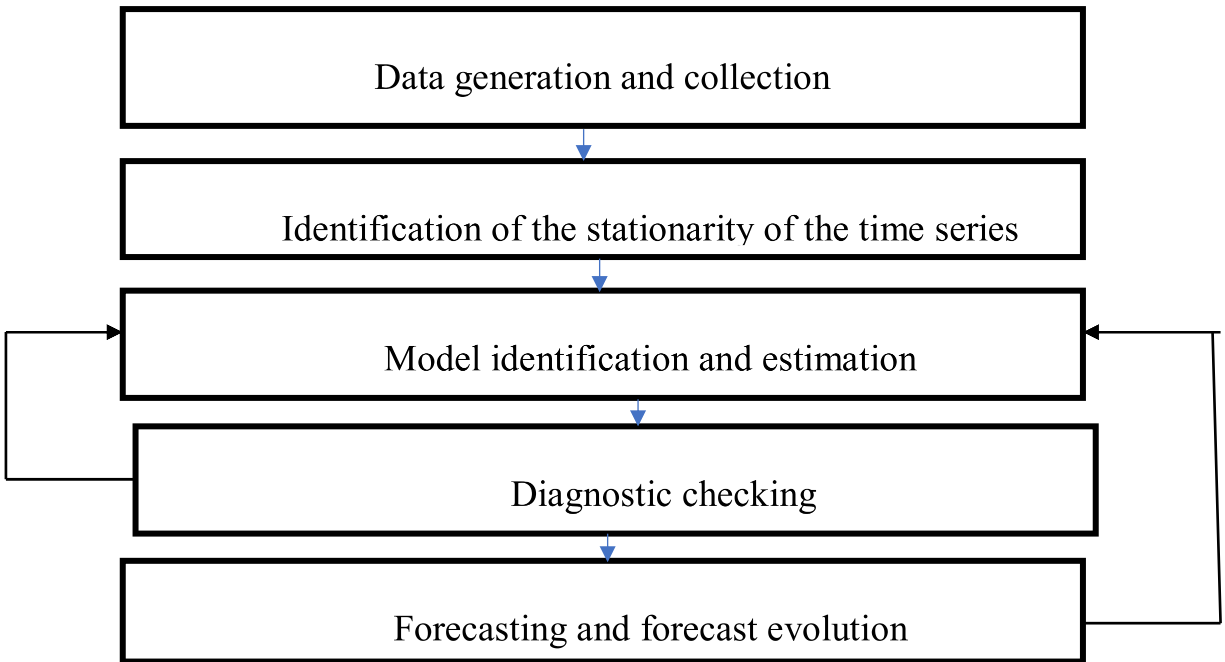
The dataset retrieved from FRED is seasonally adjusted, so it does takes into account the intra year hikes and lows and then is further divided by a GDP deflator calculated on year 2000 Prices. One important point to note is that in financial year of 2015-16, Government of India changed it GDP calculation method and hence all the GDP timer series were updated. There were many doubts regarding the veracity of new methodology[6] and hence, I have considered the data series that was compiled keeping in sync with pre 2015 GDP calculation methodology.

[3] Note : Q1, Q2, Q3 & Q4 denote - April to June, July to September, October to December and January to March quarters, respectively
[4] 1 $ = 65 Indian Rupee
[5] Organization for Economic Co-operation and Development, Gross Domestic Product by Expenditure in Constant Prices: Total Gross Domestic Product for India [NAEXKP01INQ652S], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/NAEXKP01INQ652S, May 31, 2018.
[6] Please see : https://thewire.in/economy/the-reality-of-indias-rising-gdp-numbers

4. Model and Empirical strategy

The seminal work done by Box and Jenkins (1976) led to the development of the Autoregressive Integrated Moving averages (ARIMA) models which use an iterative approach to estimate the best fit for underlying process and based on that, different approaches to forecasting. These are linear predictive models, calculated using maximum likelihood estimator in STATA, involve parameters *(p,d,q)* where *p* is the order of auto regressive terms , *d* is the order of integration or number of differences and *q* is number of moving average terms. A detailed discussion on the theoretical construct of the ARIMA model is beyond the scope of this study and in depth treatment of the models used here is discussed here in Becketti (2013).

The empirical strategy followed throughout the study is modelled along the lines of a similar exercise done for US data as illustrated in Becketti, S. (2013). Introduction To Time Series Using Stata: Modelling a real world time series: *The example of U.S. gross domestic product,* (pp. 217-270) .Texas: Stata Press. The flow chart below chalks out each step in order of analysis as undertaken in this study:



5. Modelling, Empirical Analysis and Results.

In this section I analyse the data, find an appropriate model(s) to be fit, estimate it and eventually develop a few forecasts to be compared. As it is mentioned earlier, the modelling and empirical analysis approach used throughout this dissertation relies on methodology and analysis similar to Becketti (2013, chapter 7).

5.1 Overview of the time series

The log of Indian real GDP at year 2000[7] prices as seen in figure below gives an overview of the growth rate of GDP. By regressing[8] log of real GDP series on date and annualizing slope of date we get the annual rate of growth which is 6.89%.

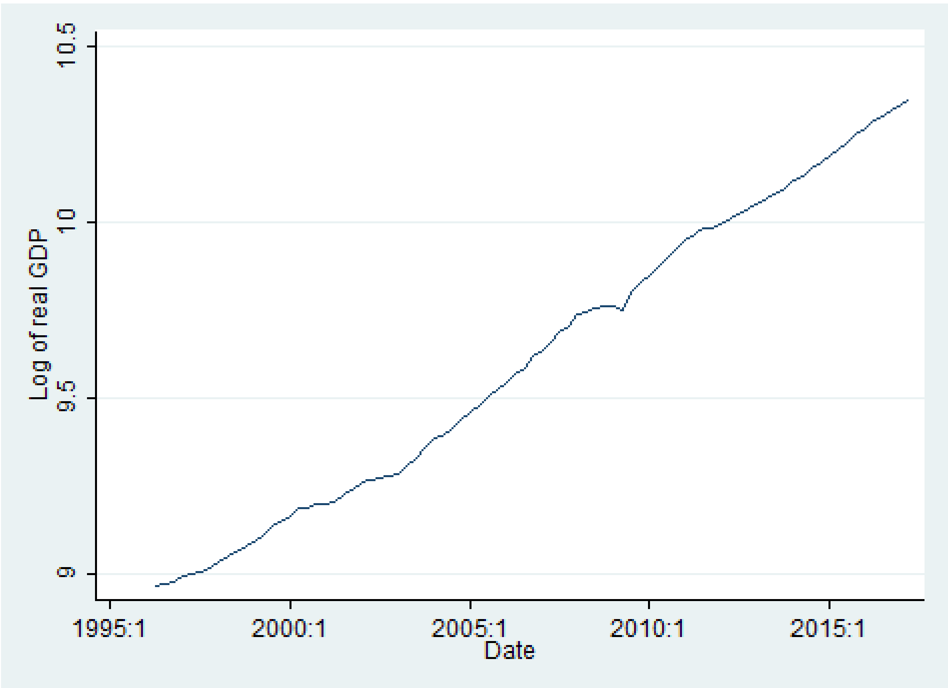


Figure 1: Log of Indian real GDP

[7] There are different datasets which use various levels of prices though FRED St. Louis, IMF and World bank have most of the chained time series process on Indian economy at 2000 price levels.
[8] Results of regression are presented in appendix.

To understand the cyclical and trend elements of the time series, Hold-Winters smoother is applied keeping in sync with the methodology adopted by Becketti (2013, chapter 7). Figure 2 below gives us two different forms the GDP time series viz linear trend and Holt-Winters smoother. We will also use Holt-Winters smooth (HWs) and trend residuals to estimate the 3 years ahead forecast. As evident from figure 2, the liner trend being fixed simply follows the annual growth percentage while HWs hovers around in factoring in local variation. Trend growth fails to account presence of any new information generation in the system and continues to mark the average rate of growth over 85 quarters as the benchmark.

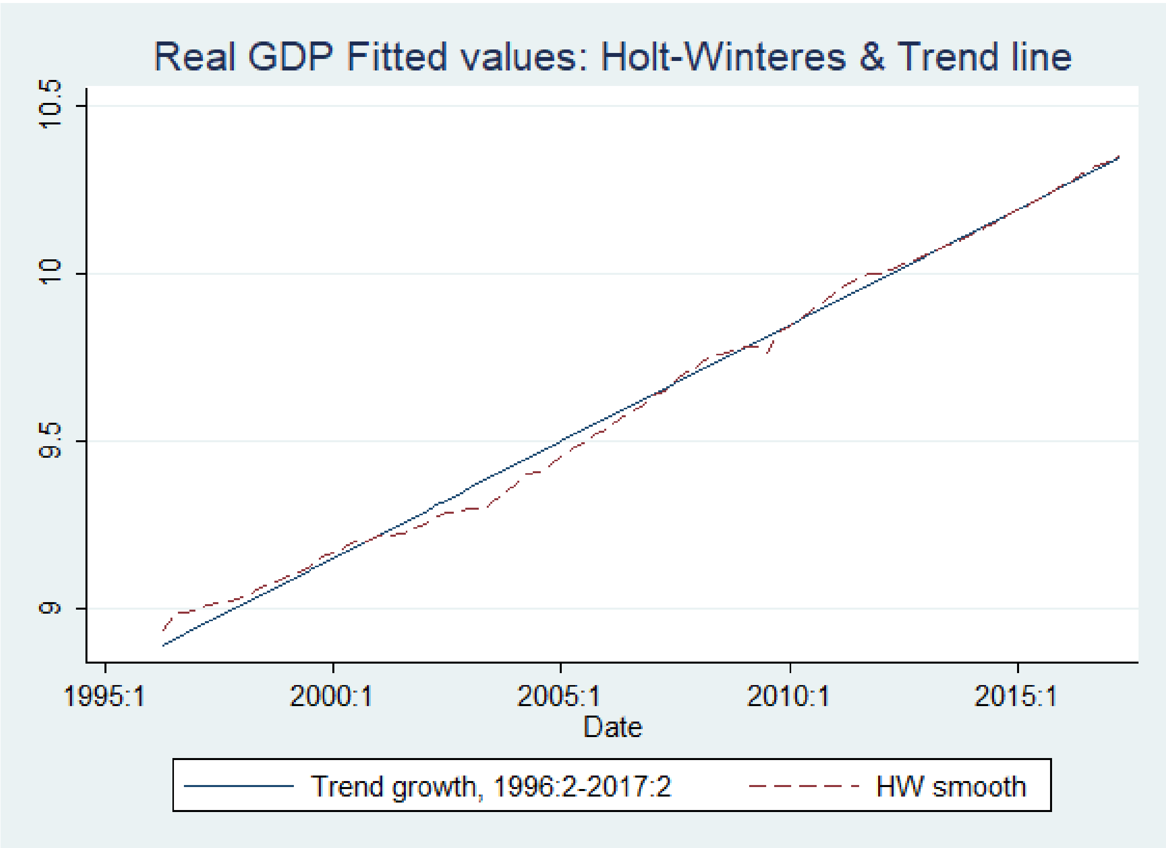


Figure 2: Two different presentation of Real GDP

In order to comprehend these to representations of real GDP time series, we shall look at the residual behaviour of both approaches. Figure 3 brings the variation in both approaches to core. As evident, the trend residuals take much larger swings across the either side of the zero marks while HWs seesaw is comparatively lesser. These continuous and prolonged swings in trend line on either side of zero are suggestive of autocorrelation. Holt-winters smoother for real GDP does not seem to rumple out the wrinkles seen in data partly showing the effect of maximum value of α as 1.

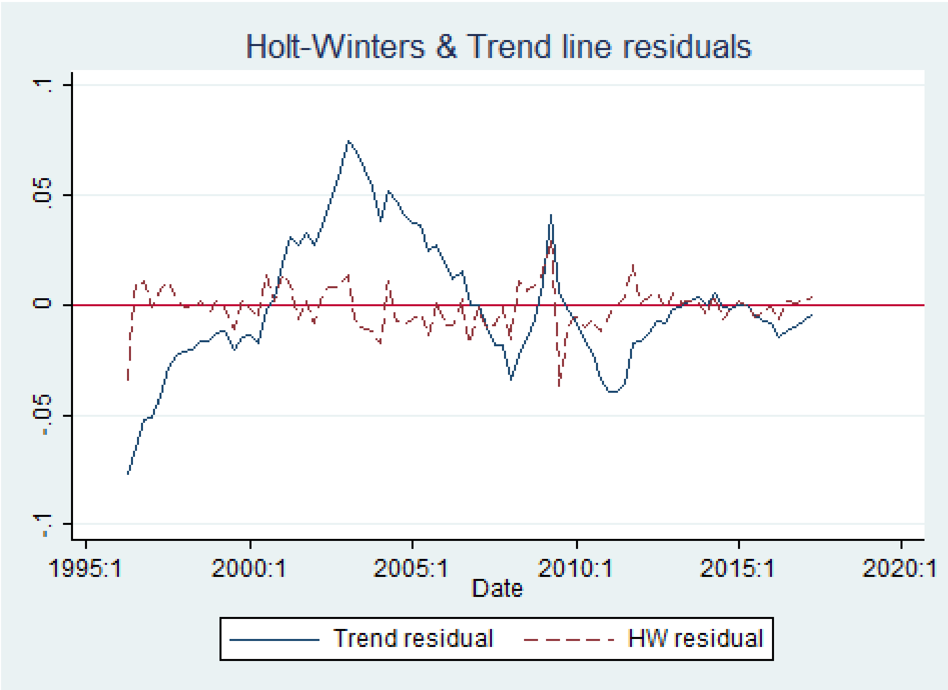


Figure 3:Holt-Winters & Trend line residuals

The two different characterization of Indian real GDP time series as discussed above if used for forecasting can shed light on underlying true model of the time series. As seen in figure 4, I have drawn the actual GDP time series along with HWs and linear trend forecast. This small exercise shows the deviation around year 2008 when financial meltdown hit the global economy. Notice how trend line forecast continue to be linearly growing ignoring any implications which might have occurred due to the 2008 meltdown but HWs does seem to take the impact of recession into consideration (as was the case in real GDP in figure 1) . The linear trend forecast seems to suggest that the slowdown in growth will not sustain and if growth was diminished for a particular time frame then it must pick up sometime in future and hence long run average would be same. Whereas, Holt-Winters smoother can be seen taking into recent events but might also be suggestive that recent phenomenon would persist meaning permanent loss of growth and shifting to an inferior steady state.

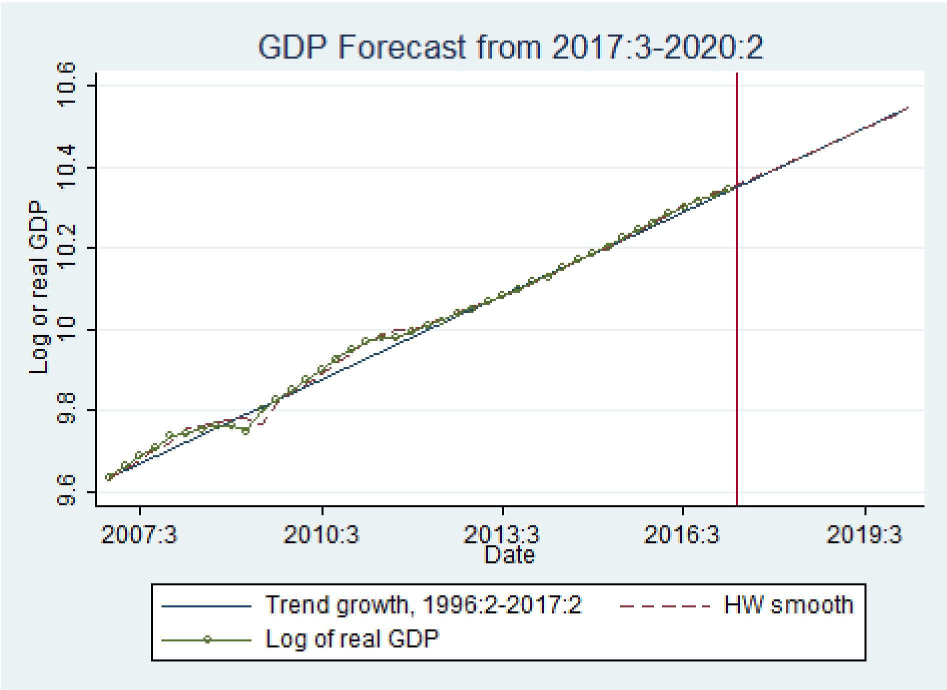


Figure 4: HW Smooth and Trend line forecasts

Both cases of the forecasts above have been very lucidly discussed in Becketti (2013, chapter 7) for United States real GDP series at 2005 prices where the impact of the global financial meltdown could be seen in a much more profound way than in Indian context. Given both the methods, HWs and linear trend, seem to suggest contradictory tales of GDP growth, perhaps combining both models to produce forecasts that will resemble one step ahead forecast will present us with a better picture. Here we will adjust the level of characterization every period as the true value is observed. Notice in the figure 5 that now both approaches yield very similar results up to 2017 and only diverge after thereafter. The linear trend process predicts GDP will grow with average growth rate as observed while HWs takes into account the recent changes.

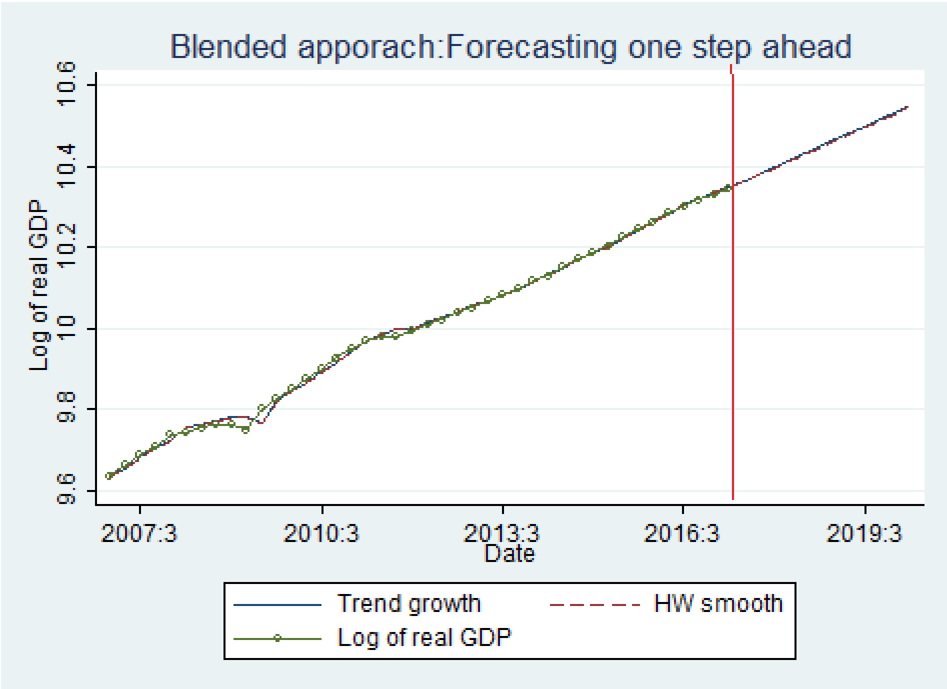


Figure 5: Blend of HWs and linear trend

The most important thing to be noted in figure 5 is that how it behaves differently than figure 4. In figure 4 where we have simply modelled the forecasting around those two approaches separately we see a straight trend line simply following the average but once we combined both transformations we see that trend and HWs act pretty similar to before forecast periods and it's only after 2017, they revert to producing the similar forecasts as in figure 4.

5.2 Model selection

Now we have seen that the aforementioned Indian real GDP series is clearly not stationary given it has an increasing trend with average growth of 6.9%. A stationarity condition would mean that long term series averages returns to a constant level which clearly is not the case for the time series in question. Earlier, the two approaches of fitting our data , linear trend and Holt-winters smoother did not yield any clear results. So here I layout the framework to fit Autoregressive moving averages (ARMA) to our data series. I proceed by checking for the stationary of the series as shown in figure 6. As we see, there is no sudden decay of auto correlation levels to zero and there is a diminishing linear trend meaning that log of real GDP series for India is not stationary.

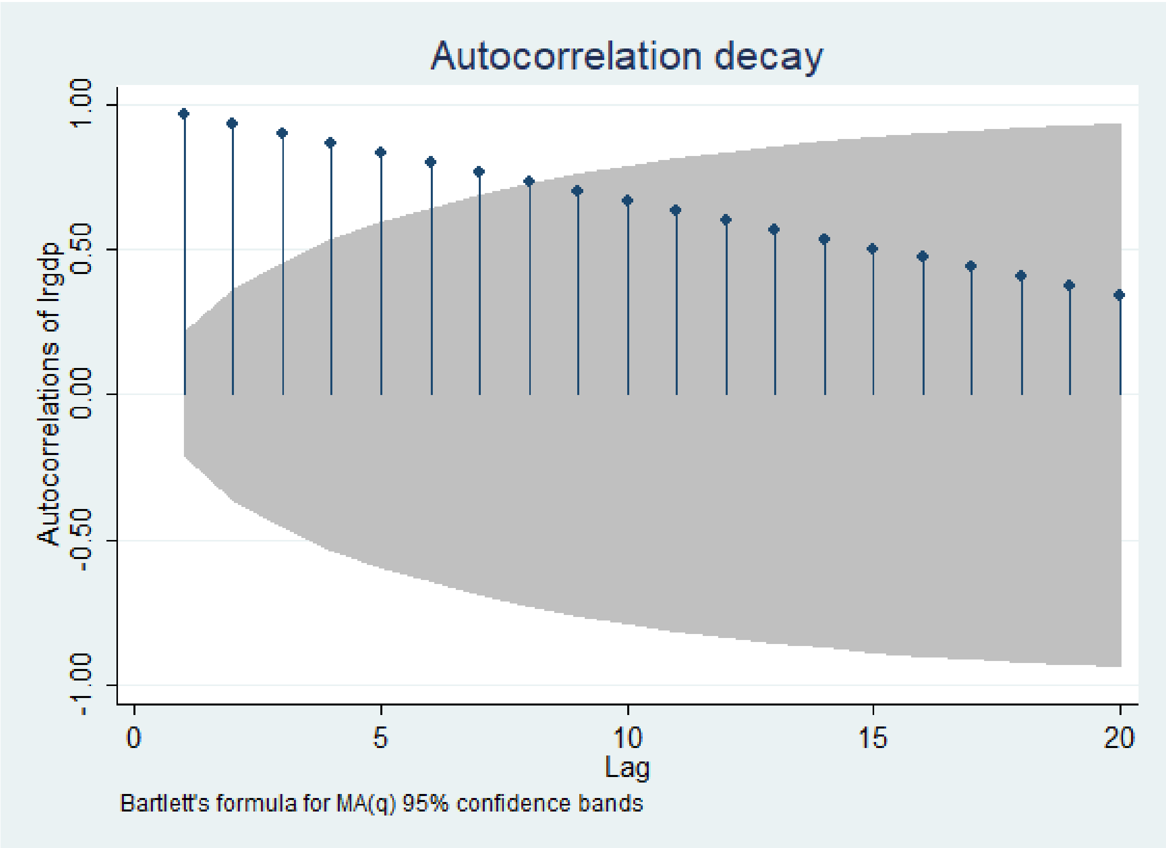


Figure 6:Autocorrelation levels of GDP

Next up, to induce stationarity, I take the first difference of the real GDP series and again calculate the auto correlations. Figure 7 clearly points out the evidence of stationarity now as autocorrelations levels quickly collapse to zero under the 95% percent confidence interval which is marked by shaded area in the figure. First three autocorrelation (AC) levels are insignificant, though they seem to be positive.

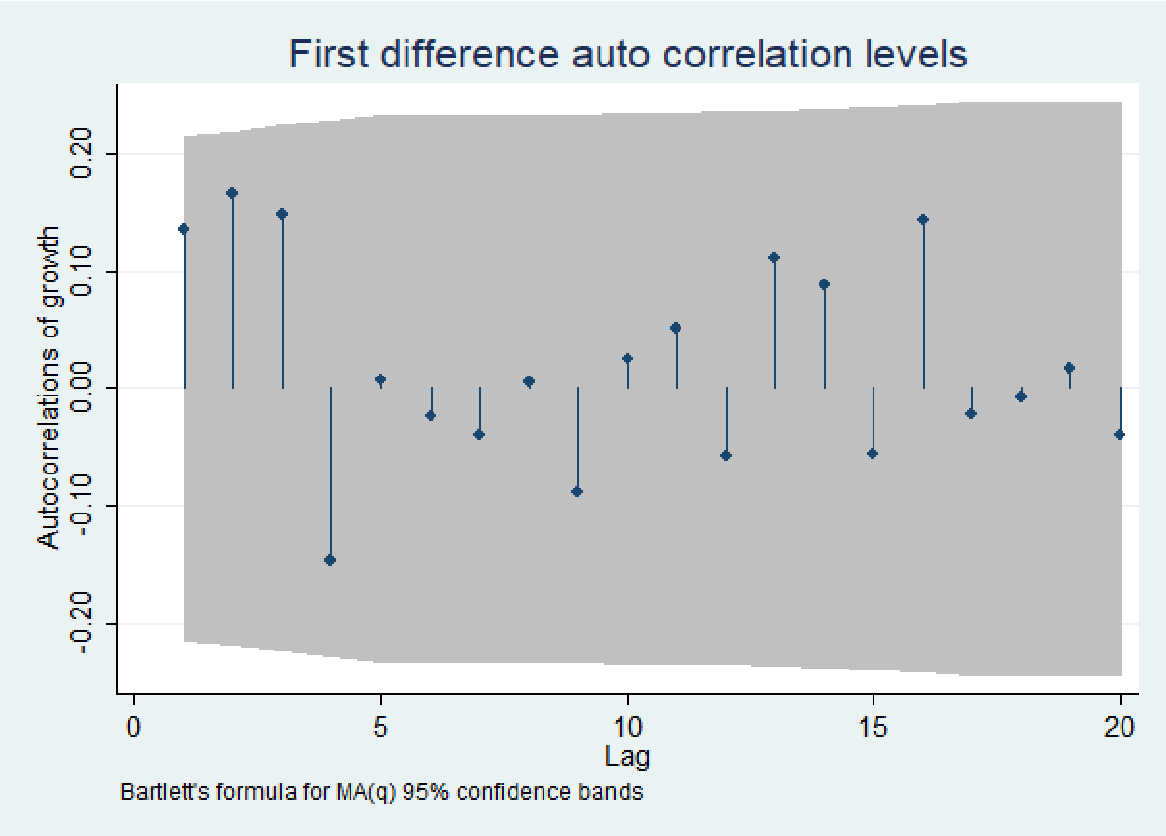


Figure 7: AC(p) levels after first difference

Now that we have established that the 1st difference of log of real is stationary, we take a step further and determine the order of auto regressive component, noted as (p) and order of moving average component labelled as (q). In the next step I calculate partial auto correlation levels. Combining graphs for AC and partial autocorrelation (PAC), we can shed some light on Ps and Qs of the time series. The figure 8 which shows AC and PAC graphs side by side provides an interesting observation. First 3 lags in both AC and PAC show marginal effects between 0.10 to 0.15, however all of them are insignificant. 4th lag of the AC and PAC shows similar effect in negative direction. Had all the lags been out of 95% confidence interval, we could have settled for the process to be ARMA(3,3). With all the lags being insignificant, there is always a chance that the log of real GDP for India is simply a white noise process meaning it cannot be forecasted. A closer inspection of the figure 7 reveals P values for autocorrelations lies in range of 0.6 to 0.7 for different lags. Formula for standard errors of partial autocorrelations is 1/√n and any effects less than 0.216 in absolute values won't be detected under 95% confidence interval. By back of the envelop calculations, data from 178 quarters would be required to be on the boundary of 95% confidence intervals.

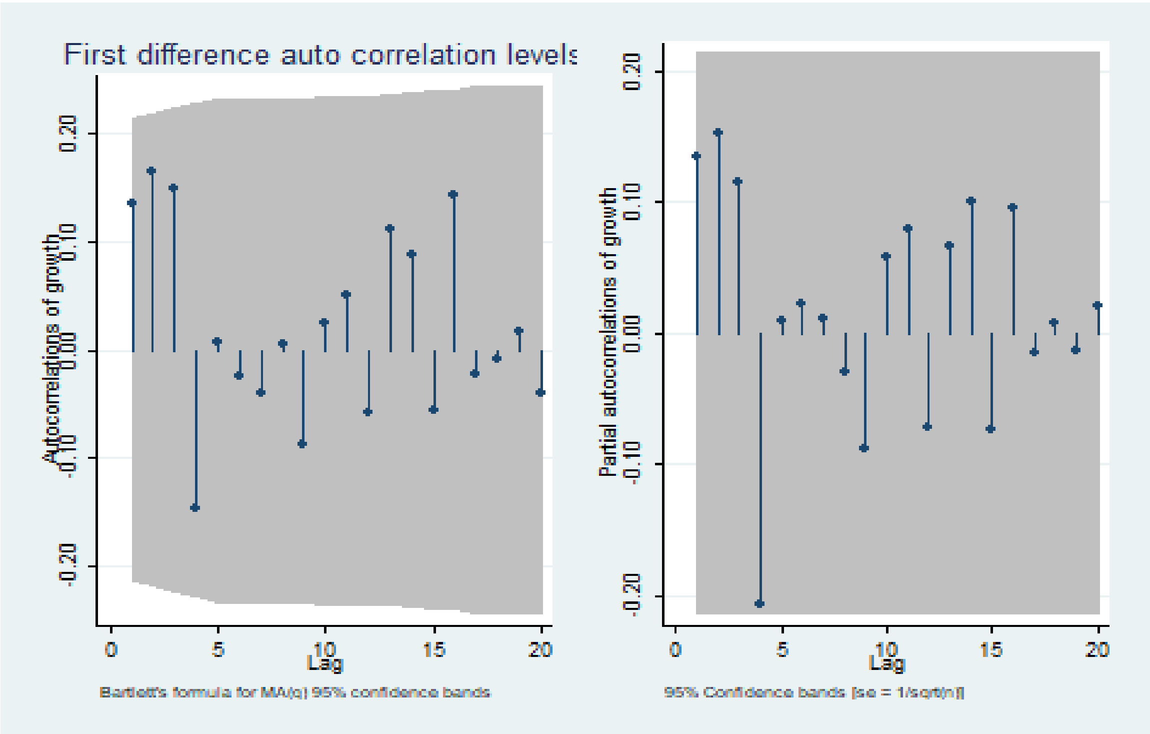


Figure 8: AC & PAC comparison

Now withstanding the apparent 'non-conclusiveness' in identifying the type of ARMA model that fits the said time series, once the data going back to 180 quarters is available, using the methodology demonstrated, we would be in a much better position to understand which ARMA model would fit the time series. For now, the possible candidates could be White Nosie, ARIMA(1,0,1) , ARIMA(1,0,2), ARIMA (2,0,2) ,ARIMA(2,0,3) and ARIMA (3,0,3). Table 2 provides a summary of all of these along with AIC and BI information criterion. Interaction of different lags of auto correlations and partial auto correlations is required to be able to judge what model fits the underlying time series. Table 1 below describes the parameters one can look at in order to conclude the same.

Table 1:Determinants of ARMA order.

Table 1: Summary of ACF & PCF with underlying process of data (real GDP 1996-2017)		
Process	*Autocorrelations functions*	*Partial ACF*
Non-stationarity	Autocorrelations(AC) do not die out, they diminish or die out linearly	
Stationarity	After 1st few lags, autocorrelations die out (collapse to 0 in form of exponential or some other combination)	
AR(p)	AC die out	PAC cut off after p lags
MA(q)	AC cut off after q lags	PAC die out
ARMA(p,q)	AC die out after 1st p-q lags	PAC die out after p-q lags

Source:Becketti (2013,p.242)

5.3 Model estimation

Table 2 details various ARIMA models and their corresponding AIC and BIC values. A glance at the table suggests that none of the models stand out.

Table 2: ARIMA regression results for various models

		ARIMA regression table				
Category	**Coefficient**	**(1,0,1)**	**(2,0,2)**	**(2,0,3)**	**(3,0,3)**	**(1,0,2)**
growth	_Constant	0.016	0.016	0.016	0.016	0.016
		(10.10)**	(11.50)**	(11.50)**	(10.13)**	(9.83)**
ARMA	L.ar	0.627	0.284	0.236	0.831	0.315
		(1.32)	(1.11)	(1.18)	(1.01)	(0.69)
	L.ma	-0.477	-0.224	-0.134	-0.721	-0.227
		(0.90)	(0.89)	(0.00)	(0.01)	(0.47)
	L2.ar		-0.478	-0.765	-0.940	
			(1.81)	(5.91)**	(3.32)**	
	L2.ma		0.708	0.968	1.120	0.222
			(3.43)**	(0.00)	(0.00)	(1.47)
	L3.ma			0.124	-0.462	
				(0.00)	(0.00)	
	L3.ar				0.466	
					(0.84)	
sigma	_cons	0.009	0.009	0.009	0.009	0.009
		(17.22)**	(15.85)**	(0.00)	(0.00)	(17.56)**
N		84	84	84	84	84

I will proceed further keeping in the mind the other objective of this study which is to demonstrate a comprehensive empirical methodology for fitting models in context of

Indian GDP data to understand which ARMA model would fit the time series we shall follow Becketti (2013, pp. 217-269) empirical strategy and select ARMA(1,1) and ARMA (1,2) and proceed with model estimation.

Table 3: Information Criterion

	ARIMA models				
Information Criterion	**(1,0,1)**	**(2,0,2)**	**(2,0,3)**	**(3,0,3)**	**(1,0,2)**
AIC	-541.6528	-541.1647	-540.7496	-539.2596	-541.5216
BIC	-531.9296	-526.5798	-523.7339	-519.8131	-529.3675

As evident in table 2 there is no certainty that which ARIMA models fits best to the log real GDP. To be able to compare the different ARIMA models, one approach is to compare the coefficients estimates of underlying stochastics process: $y_t = \psi(L)\epsilon_t$.

Table 4: psi weights for ARIMA models

ARIMA Models		
Model	**(1,0,2)**	**(1,0,1)**
psi1	0.542	1.105
Psi2	-0.052	0.693
Psi3	-0.016	0.435
Psi4	-0.005	0.273

Results from table 4 suggest that ARIMA (1,0,2) effects are persistent during 1^{st} lag and a change in sign from positive to negative but quickly die out where as ARIMA (1,0,1) reveals that effects persist over a long time before eventually decaying. It is still inconclusive that which model might be a better representation of the data.

5.4 Model Diagnostics

As discussed in previous section, there is a case to be made that many models yield similar results though none of them describing the exact data series. I perform a Q-test and create a cumulative periodogram of residual in order to see if there is any indication of white noise in the data. Q-test does not provide any proof that residuals deviate from white noise and looking at cumulative periodogram it is clear that residual do not systematically take excursions away form 45-degree line and stay north of confidence intervals.

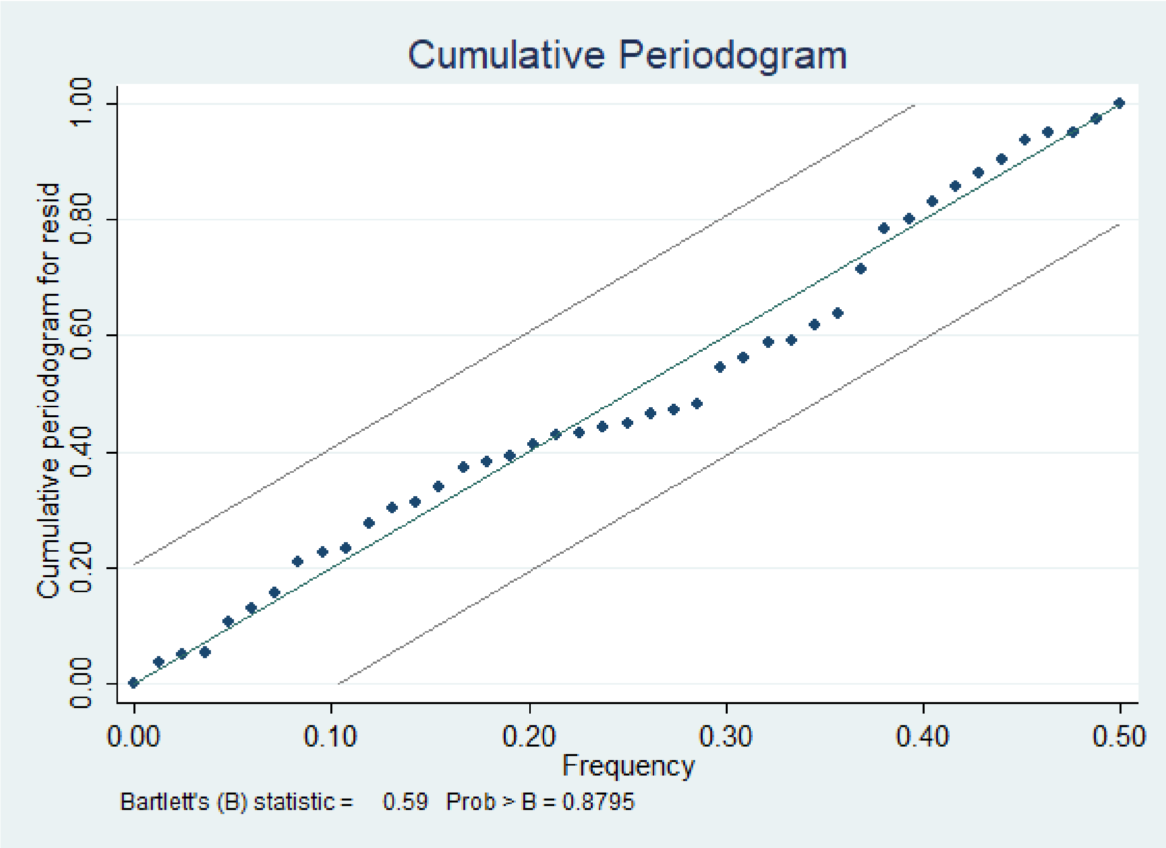


Figure 9: Cumulative Periodogram

5.4 Forecasting

Now that we have estimated[9] the ARMA (1,2) and ARMA(1,1) models, I calculate four difference forecast for each of the model. Mainly, four types of forecasts[10] viz one step ahead, structural, dynamic and time constant is used.

[9] Model estimation results in appendix.
[10] For further discussion on forecasts, see Becketti S. (2013) Modelling a real-world time series. *Introduction to time series using STATA* . pp (257-261)

The table below details forecast for ARMA (1,2) using different approaches[11] for initial five years, that from 1996 to 2000. The table 6 below underlines interesting patterns.

Table 5: Forecast evolution for first 5 quarters in sample for AR (1)

Date	agrowth	arxb	arst	ardyYR	art0
1996:2	.	6.5	6.5	6.5	.
1996:3	2.2	6.5	6.5	6.5	.
1996:4	1.9	5.9	6.5	5.9	.
1997:1	6.3	5.9	6.5	5.9	.
1997:2	3.0	6.5	6.5	6.5	.

The *agrowth* column is the first differenced growth rate of log of real GDP and hence the missing value, *axrb* is one step ahead forecast , *arst* is simply the structural forecast showing the mean of the underlying process, *ardyYR* is the dynamic forecast which which would come into effect by 2008 and for now just stays true to the *arxb.* The art0 forecast would come into effect by 2008 as well. Similarly, table 7 below shows various estimates of the forcasts for MA(2) model for first five quarters.

Table 6: Forecast evolution for first 5 quaters in sample for MA(2)

Date	agrowth	maxb	mast	mady	art0
1996:2	.	6.5	6.5	6.5	.
1996:3	2.2	6.5	6.5	6.5	.
1996:4	1.9	6.3	6.5	6.3	.
1997:1	6.3	5.4	6.5	5.4	.
1997:2	3.0	5.6	6.5	5.6	.

The graph below highlights the forecast evolution[12] after 2008 for MA(2) model. It can be seen that *maxb* and *mady* closely predict each other upto 2006 but after that both diverge and once again converge after 2008. Structural forecast remains true to the mean

[11] Forecast evolution for all years in sample is listed as a table in appendix.
[12] Table listing detailed residuals and forecast estimates for all 85 quarters availabl in appendix.

other process .Forecast for *art0* starts after 2008 and show less volatility than other three forecasts.

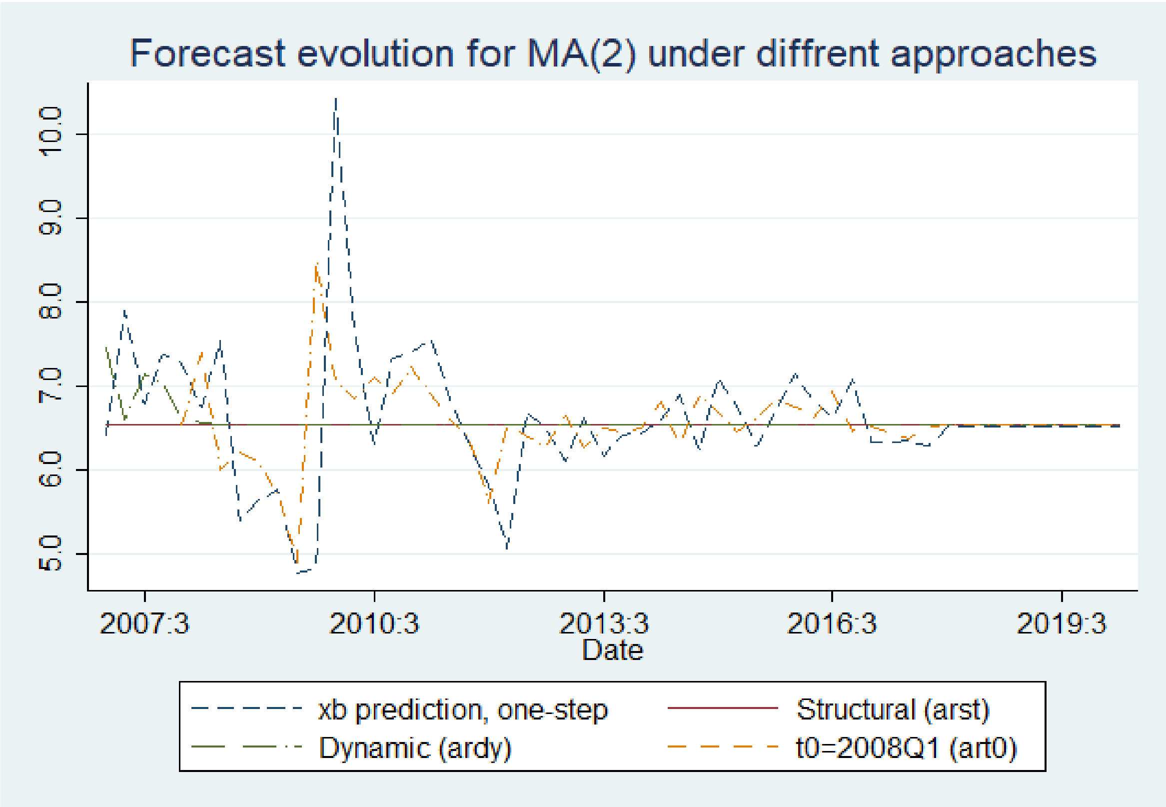


Figure 10: Forecast evolution for MA(2)

5.5 Forecast Evolution

One of the ways to see how forecast plays out under different models is to compare the actual growth rates with mean of forecast estimates produced by all the models we considered for out of sample (1996:2 to 2012:2) and in-sample periods (2012:3 to 2017:2).

Table 7: Growth rate evolution

Period	Mean(growthYR)	Mean(ardg)	Mean(madg)	Mean(hwg)	Mean(tg)
In sample	6.5	6.5	6.5	6.7	6.8
Out of sample	6.7	6.5	6.5	6.7	6.8
Total	6.6	6.5	6.5	6.7	6.8

Table above show us the mean forecast for in-sample and out-sample periods. The actual mean of growth between sample periods is 6.5 and 6.7% but none of the forecasts predict that. However, in order to ascertain the most efficient model specification and see the deviation of residuals, table 8 shows that in-sample residual variance is much higher than out of smaple. ARIMA models have lower residual variance in comparison to the linear trend and holt-winters approach.

Table 8: Standard deviations for redisdual of different forecast models

Period	sd(ardgres)	sd(madgres)	sd(hwgres)	sd(hwg)
In sample	4.2	4.2	4.6	0.0
Out of sample	1.5	1.5	1.5	0.0
Total	3.7	3.7	4.1	0.0

The figure below shows forecast evolution for with-in sample and out of sample data. HW smooth keeps close to real GDP by following it one period ahead. ARMA models show a

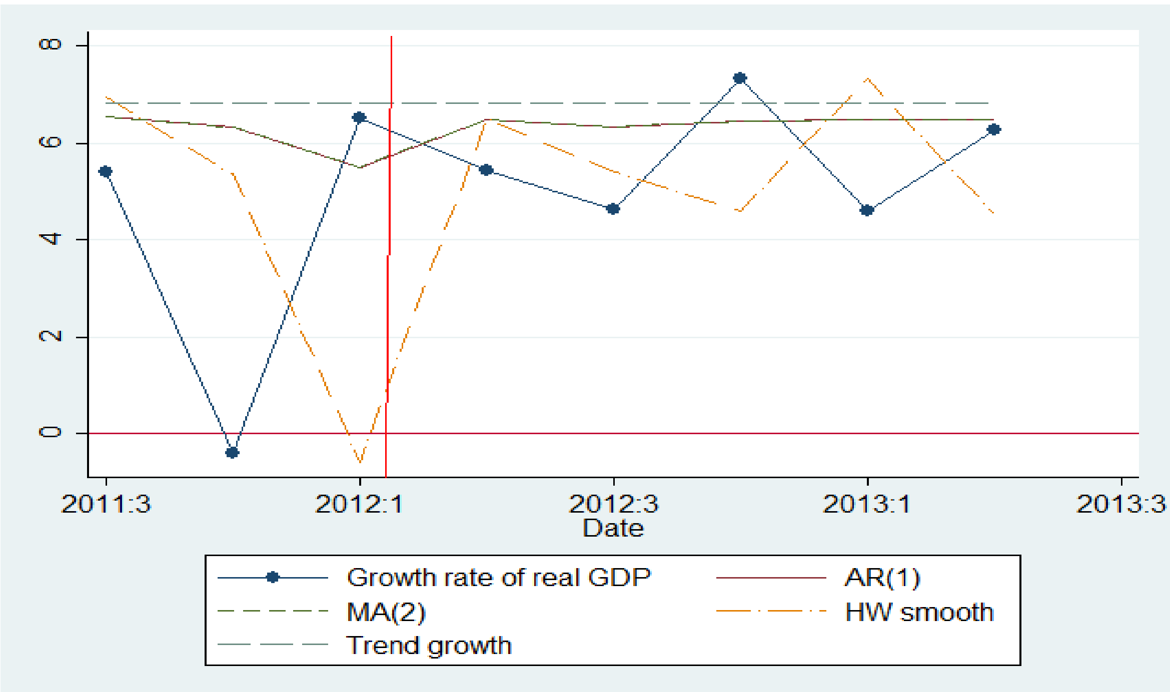


Figure 11: Forecast evolution around 2012

Muted response than HW and ultimately, all forecast seem to go back to average levels as the out-of-sample data range comes to an end.

Figure below gives an overview of all the forecasting approaches from 2008 to 2017 period of study, marking the impact of 2008 financial crisis.

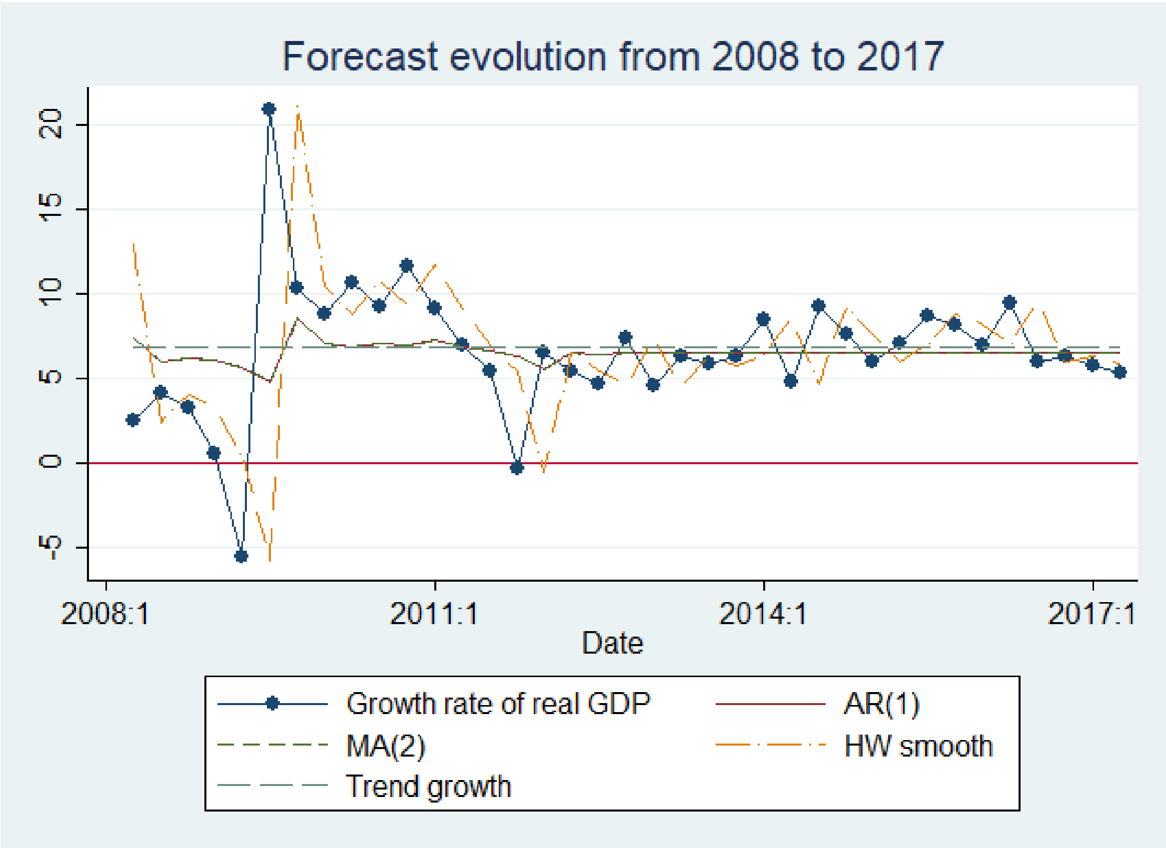


Figure 12: Forecast evolution for after 2008.

7. Conclusion

This study is aimed at modelling Indian real GDP growth time series data from 1996 to 2017, about 85 quarters. After looking at the initial time series data, I fit HW smooth and linear trend to the time series and note both forecast model follows each other closely. Then in order to estimate an ARMA mode, I induced stationarity and see that AR levels do not suddenly collapse to zero. Further proceeding with modelling estimation, I find that no ARIMA specification in particular seemed better over any other or yielding any significant result. In order to attain the secondary goal of the study of demonstrating a comprehensive approach to the forecasting the GDP series for developing nations, I select AR(1) and MA (2) specifications . Subsequently, I calculate dynamic, structural and fixed time (2008) forecast and find that long run forecasts tend to converge for Indian GDP .

One of the most glaring limitation of this study remains the limited amount of quarterly data available for real GDP series for India. Due to the limited data, the auto correlation level values remain insignificant and same is the case with model selection being non-conclusive in terms of fit to the underlying data generation process. A similar study for done for USA as demonstrated in Becketti (2013) is able to generate much more pronounced effects given the data comprised of 260 quarters.

Now withstanding the lack of ample data for analysis, I believe that analysis is necessary in order to ascertain any significant effect and models of variance analysis such as ARCH-GARCH shall be considered. Further, instead of just looking at the log of real GDP, one can also foray into construing a macroeconomy model using VECM and VAR to understand the dynamics of economy.

References

Becketti, S. (2013). Introduction To Time Series Using Stata: Modelling a real world time series: *The example of U.S. gross domestic product,* (pp. 217-270). Texas: Stata Press

Becketti, S. (2013). Introduction to Time Series Using Stata.Texas: Stata Press

Box, George E. P. and Gwilym M. Jenkins (1976). *Time Series Analysis: Forecasting and Control*, Revised Edition, Oakland, CA: Holden-Day.

Das, J. and Do, Q (2014). *US and them: The geography of academic research,* VOX CEPR's policy portal. Retrieved from: https://voxeu.org/article/geographical-bias-top-journal-publication.

Musundi Sammy et al (2016). Modelling and Forecasting Kenyan GDP Using Autoregressive Integrated Moving Average (ARIMA) Models. *Science Journal of Applied Mathematics and Statistics*. Vol. 4, No. 2, 2016, pp. 64-73. doi:10.11648/j.sjams.20160402.18

Maity, B., & Chatterjee, B. (2012). Forecasting GDP growth rates of India: An empirical study. *International Journal of Economics and Management Sciences*, 1(9), 52-58.

Changle, R. & Matharu, S. (n.d.). A Study Forecasting the GDP of India Using ARIMA Model.

Dristaki, C. (2015). Forecasting Real GDP Rate through Econometric Models: An Empirical Study from Greece. *Journal of International Business and Economics*
June 2015, Vol. 3, No. 1, pp. 13-19

Zakai, M. (2014). A time series modeling on GDP of Pakistan. *Journal of Contemporary Issues in Business research.* 3(4), 200-210.

Zhang, H. (2013). Modeling and forecasting regional GDP in Sweden using autoregressive models. *Working Paper*, Högskolan Dalarna University, Sweden

Appendix

1.Regression results:

	lrgdp
DATE	0.017
	(129.70)**
_cons	6.392
	(254.86)**
R^2	1.00
N	85

* p<0.05; ** p<0.01

2.Model diagnostics results

growth	_cons	0.016	0.016	0.016
		(10.44)**	(8.87)**	(11.20)**
ARMA	L.ar	0.117		
		(1.31)		
	L2.ar	0.154		
		(1.25)		
ARMA	L.ma		0.152	0.053
			(1.52)	(0.68)
	L2.ma		0.225	0.224
			(1.45)	(1.89)
	L3.ma		0.198	
			(1.31)	
sigma	_cons	0.009	0.009	0.009
		(17.25)**	(17.40)**	(17.13)**
N		84	84	84

* p<0.05; ** p<0.01

3.Forecasting with AR (1)

growth	_cons	0.016
		(12.17)**
ARMA	L.ar	0.136
		(1.60)
sigma	_cons	0.009
		(16.97)**
N		84

* p<0.05; ** p<0.01

4.Regressing log of real GDP on DATE after 2012 Q3

	lrgdp
DATE	0.017
	(75.64)**
_cons	6.419
	(159.66)**
R^2	0.99
N	65

* p<0.05; ** p<0.01

5.Regressing yearly growth on AR(1) and ARMA(2) after 2012 Q3

growthYR	_cons	6.470	6.470
		(8.56)**	(8.56)**
ARMA	L.ma	0.066	0.066
		(0.64)	(0.64)
	L2.ma	0.215	0.215
		(1.40)	(1.40)
sigma	_cons	4.107	4.107
		(12.98)**	(12.98)**
N		64	64

* p<0.05; ** p<0.01

Table illustrating *means & residuals* of different models around 2012 Q2

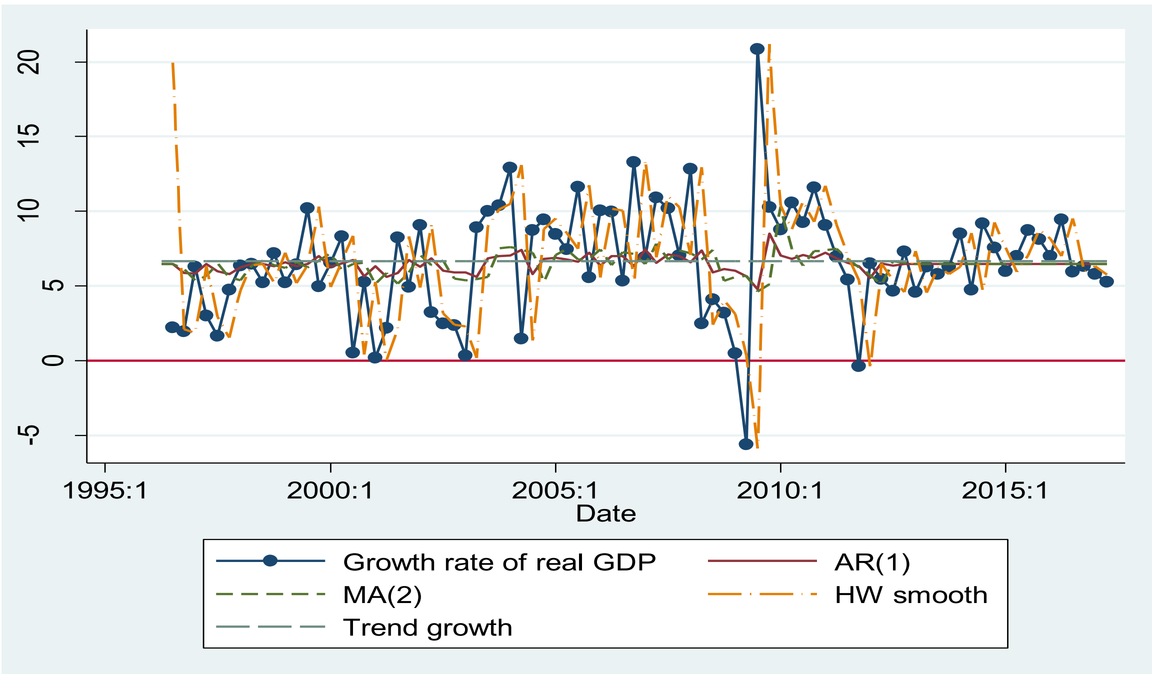
Date	mean(grow~R)	mean(ardg)	mean(madg)	mean(hwg)	mean(tg)
2012:3	4.6	6.3	6.7	5.4	6.7
2012:4	7.3	6.5	6.5	4.6	6.7
2013:1	4.6	6.5	6.5	7.3	6.7
2013:2	6.3	6.5	6.5	4.5	6.7
Total	5.7	6.4	6.5	5.5	6.7

period	mean(ardgres)	mean(madgres)	mean(hwgres)	mean(tgres)
In sample	-0.0	-0.0	6.7	6.8
Out of sample	-0.2	-0.2	6.7	6.8
Total	-0.1	-0.1	6.7	6.8

Table illustrating *means* of residuals of different models

Date	mean(ardgres)	mean(madgres)	mean(hwgres)	mean(tgres)
2012:3	0.0	2.0	5.4	6.6
2012:4	-0.0	-0.9	4.6	6.6
2013:1	0.0	1.5	7.3	6.6
2013:2	0.0	0.3	4.5	6.6
Total	0.0	0.7	5.5	6.6

7.Entire sample forecast evolution compared with actual growth rates:



The figure below highlights the forecast evolution after 2008 for AR(1) model: