Analysis of Meat Production in the United States

Group 1 – Jocelyn Hunyadi, Elizabeth Schneider, Jack Kramer, Karleshia Dorsey, & Ken Marciel

1

Project Objectives

Primary Goal – Investigate time series data of meat production in the US

- Have significant epidemiological events (i.e. pandemics, epidemics) influenced trends in meat production?
- Have historical and economic events (i.e. recession) influenced trends in meat production?

Secondary Goal – Analyze and compare trends in production between total red meat and total poultry production (federally inspected)

2

Data Overview

Time: Monthly data from January 1983 - April 2021.

Meat Types:

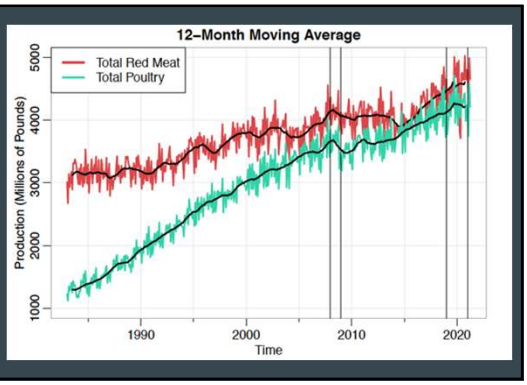
1. Red Meat – Beef, Veal, Lamb & Mutton, and Pork
2. Poultry – Broilers, Other Chicken, and Turkey

Production: Response variable - units: millions of pounds

Additional Variables: Time (months), Population (thousands), Median Income (monthly)

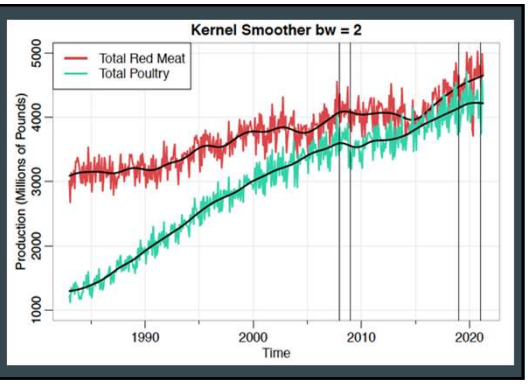
3

Red Meat & Poultry: 12-month moving average



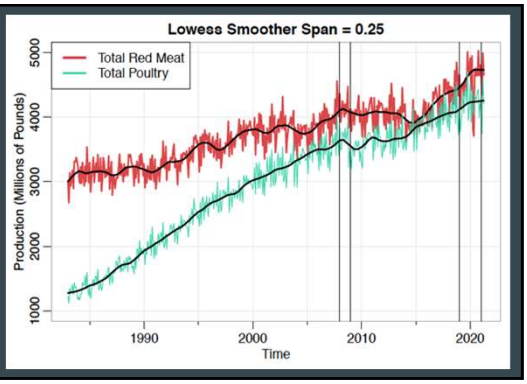
4

Red Meat & Poultry: Kernel Smoother



5

Red Meat & Poultry: Lowess Smoother



6

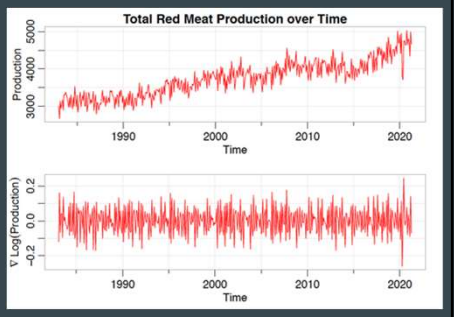
Analysis of Red Meat Production – ARIMA Model

7

Transformation to Stationarity

Differencing the log production appears to improve the time series to stationarity.

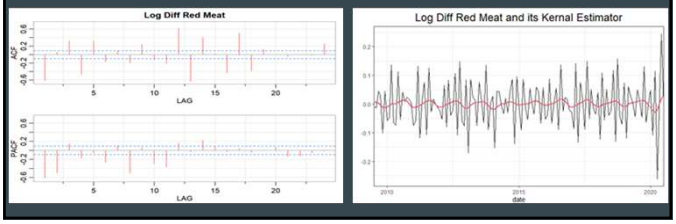
DF, ADF, and PP test all had p-value < 0.01



8

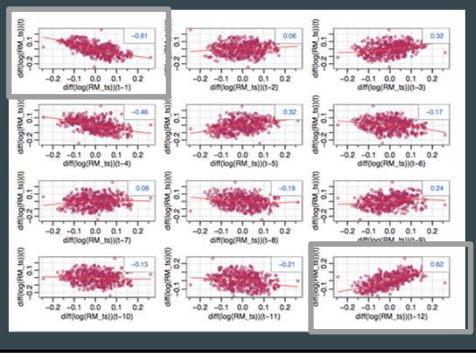
Transformation to Stationarity

Some seasonal pattern still exists after differencing, thus a SARIMA model should be considered



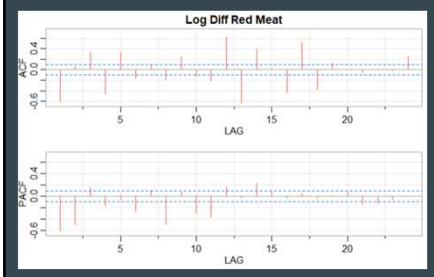
9

Red Meat: Transformed



10

Preliminary TS Model Selection



ARIMA(12,1,0) model (though complex) may be appropriate for red meat based on the PACF.

11

Model 1 - ARIMA(12,1,0)

Parameter estimates for Yule-Walker and OLS were approximately equivalent.

$M_t = -0.75_{(0.047)}M_{t-1} - 0.61_{(0.058)}M_{t-2} - \ldots + 0.16_{(0.047)}M_{t-12}$

...where M_t is total red meat production.

AIC: -3.695

AICc:-3.694

BIC:-3.569

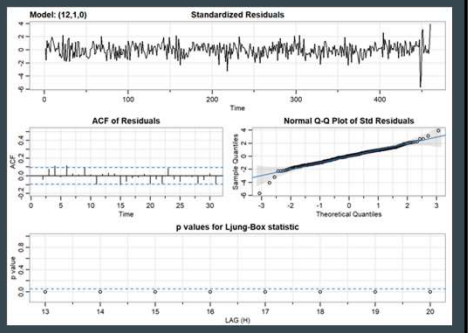
AR(12) Model for Red Meat Production

YWE	OLS
$\hat{\varphi}_1 = -0.7526_{(0.0467)}$,	$\hat{\varphi}_1 = -0.7501_{(0.0473)}$,
$\hat{\varphi}_2 = -0.6148_{(0.0577)}$,	$\hat{\varphi}_2 = -0.6282_{(0.0585)}$,
$\hat{\varphi}_3 = -0.3196_{(0.0611)}$,	$\hat{\varphi}_3 = -0.3297_{(0.0619)}$,
$\hat{\varphi}_4 = -0.6084_{(0.0610)}$,	$\hat{\varphi}_4 = -0.6168_{(0.0615)}$,
$\hat{\varphi}_5 = -0.3738_{(0.0624)}$,	$\hat{\varphi}_5 = -0.3856_{(0.0626)}$,
$\hat{\varphi}_6 = -0.5343_{(0.0614)}$,	$\hat{\varphi}_6 = -0.5511_{(0.0616)}$,
$\hat{\varphi}_7 = -0.4375_{(0.0614)}$,	$\hat{\varphi}_7 = -0.4463_{(0.0622)}$,
$\hat{\varphi}_8 = -0.5449_{(0.0624)}$,	$\hat{\varphi}_8 = -0.5388_{(0.0632)}$,
$\hat{\varphi}_9 = -0.3238_{(0.0610)}$,	$\hat{\varphi}_9 = -0.3244_{(0.0610)}$,
$\hat{\varphi}_{10} = -0.4440_{(0.0611)}$,	$\hat{\varphi}_{10} = -0.4552_{(0.0620)}$,
$\hat{\varphi}_{11} = -0.2403_{(0.0577)}$,	$\hat{\varphi}_{11} = -0.2477_{(0.0604)}$,
$\hat{\varphi}_{12} = 0.1599_{(0.0467)}$	$\hat{\varphi}_{12} = 0.1738_{(0.0489)}$

12

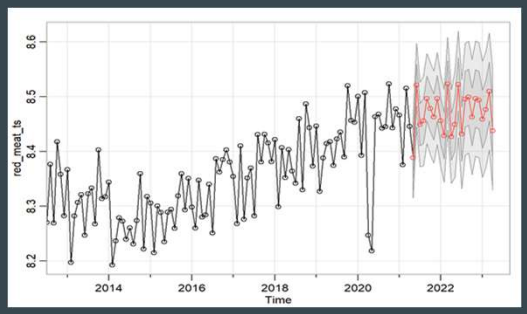
Model 1 – Diagnostics

- Spike in the residuals plot in 2020 (COVID-19)
- Significant p-values for Q-stat



13

Model 1 – Forecast



14

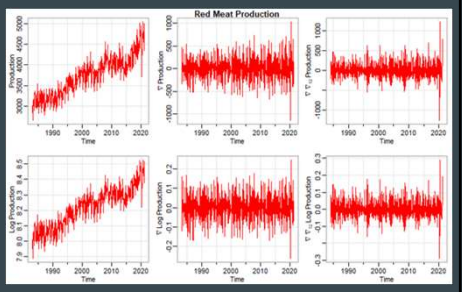
Analysis of Red Meat Production – SARIMA Model

15

Model 2 - SARIMA

The trend is removed by differencing, but the transformations do not remove the spike in variance in 2020

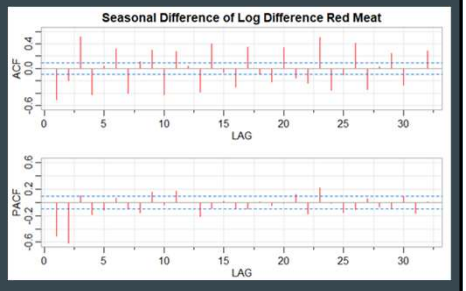
Possible indication of COVID impact on meat production?



16

Model 2 - SARIMA

The PACF appears to cut off after lag 2.



17

Model 2 - Model Selection Process

Based on AIC, AICc, and BIC values, the $SARIMA(3,1,3)x(1,1,1)_{12}$ is the best seasonal model choice.

p	d	q	P	D	Q	S	DOF	AIC	AICc	BIC
0	1	0	1	1	0	12	446	-2.53752	-2.5375	-2.5196
1	1	0	1	1	0	12	445	-2.82721	-2.82715	-2.80034
2	1	0	1	1	0	12	444	-3.36642	-3.36631	-3.33059
3	1	0	1	1	0	12	443	-3.36577	-3.36557	-3.32098
4	1	0	1	1	0	12	442	-3.4008	-3.40051	-3.34706
0	1	0	1	1	0	12	446	-2.53752	-2.5375	-2.5196
0	1	1	0	0	0	12	445	-3.1483	-3.14824	-3.12143
0	1	2	1	0	0	12	444	-3.19585	-3.19574	-3.16002
0	1	3	1	0	0	12	443	-3.21393	-3.21374	-3.16914
0	1	4	1	0	0	12	442	-3.34901	-3.34872	-3.29526
0	1	0	1	0	0	12	446	-2.53752	-2.5375	-2.5196
1	1	1	1	0	0	12	444	-3.17419	-3.17408	-3.13836
2	1	2	1	0	0	12	442	-3.39004	-3.38975	-3.33629
3	1	3	1	0	0	12	440	-3.43341	-3.43287	-3.36175
0	1	0	1	0	1	12	436	-3.61918	-3.61789	-3.51169
1	1	1	1	0	1	12	436	-3.61918	-3.61789	-3.51169
2	1	2	1	0	1	12	436	-3.61918	-3.61789	-3.51169
3	1	3	1	1	1	12	439	-3.91218	-3.91148	-3.83156

18

Slide 14

1 my forecast plot is different than this one, will update based on new one R generates
Jack Kramer, 8/1/2021

1 I believe this was from Ken's original forecast file in his folder. The jpg for it are there
Elizabeth Schneider, 8/1/2021

2 I think the difference then is using d=1 to d=0. This plot looks like it differences in the data so d=0. I think Sarima.for function will convert y axis to better label if you use d=1
Jack Kramer, 8/1/2021

1 Update Plot with cleaned y-axis label
Jocelyn Hunyadi, 8/1/2021

3 I put my forecast plot below, I couldnt get sarima.for to change the ylabel
Jack Kramer, 8/1/2021

Model 2 - SARIMA(3,1,3)x(1,1,1)12

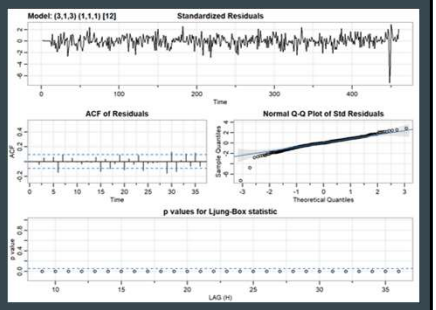
```
Parameters for SARMA(3,1,3)(1,1,1)[12] model

##      Estimate     SE  t.value p.value
## ar1   -1.0951 0.0675 -16.2301  0.0000
## ar2   -0.9267 0.0782 -11.8578  0.0000
## ar3    0.0632 0.0674   0.9371  0.3492
## ma1    0.3788 0.0468   8.0900  0.0000
## ma2    0.1141 0.0519   2.1999  0.0283
## ma3   -0.7707 0.0438 -17.5921  0.0000
## sar1   0.1366 0.0550   2.4847  0.0133
## sma1  -0.9225 0.0278 -33.2408  0.0000
```

Note that the AR3 parameter is **not** significant.

19

Model 2 - Diagnostics

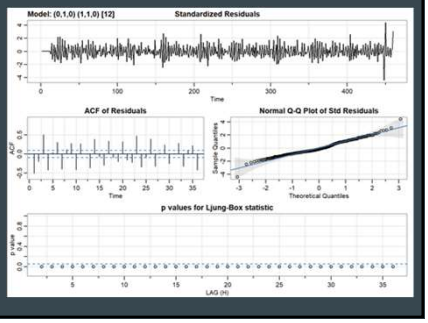
- Note large spike in residual plot.
- ACF of residuals mostly not significant
- Q-stats, however, are all significant indicating that the residuals are not white noise.



20

Model 2 - Comparison

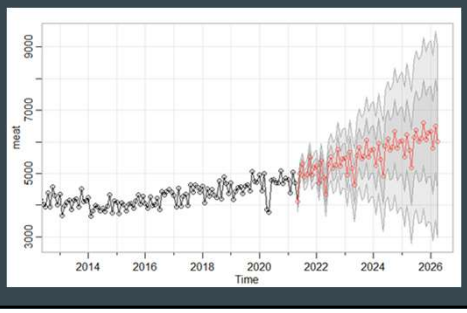
- SARIMA$(0,1,0)x(1,1,0)_{12}$
- Notice multiple large and significant correlations in the ACF
- Q-stat values are all significant.



21

Model 2 - Forecasts

Forecasts predict a slight increase in red meat total production (on average) over the next 2 years.



22

May, June and July Actual vs Forecast

Total Red Meat in Millions of Pounds

	Actual	ARIMA(12,1,0)	SARIMA(3,1,3)(1,1,1)12
May	4246.3	4397.2	4448.9
June	4617.8	5019.5	4941.2

23

Analysis of Red Meat Production – Regression with AutoCorrelated Errors

24

Slide 22

4 I put my same forecast model on the right, Once again i couldnt change ylabel and also not sure how this one was transformed out of Log scale. EIther one works for me, however, we cant compare this forecast plot with other model if we keep this one with log removed
Jack Kramer, 8/2/2021

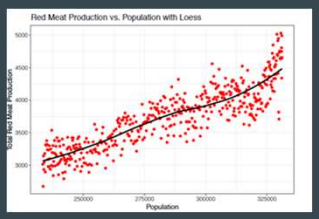
1 I fed the untransformed data into the sarima.for function. It doesn't let you feed in the linear model like the sarima function does.
Ken Marciel, 8/2/2021

Slide 23

5 I added the actuals for red meat production and the forecasted values for ARIMA and SARIMA models. I dont have the thrid model, so can someone else load in those forecasts?
Jack Kramer, 8/2/2021

6 i think this is a cool talking point for the presentation - probly dont hav to incorporate into the paper
Jack Kramer, 8/2/2021

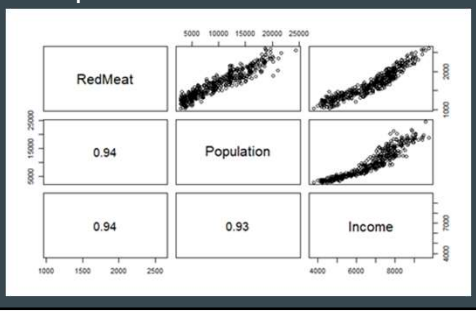
Covariates – Population & Personal Income



Covariates generally have a linear relationship with Red Meat Production with slight underlying curvature as indicated by the Loess smoother.

25

Covariates – Population & Personal Income



26

Model 3 - Initial Linear Regression

Model	k	SSE	df	MSE	R2	AIC	BIC
M1	2	23268197	458	50804	0.7572	11.8444	13.7092
M2	3	23265587	457	50909	0.7572	11.8487	13.7225
M3	4	23223123	456	50928	0.7577	11.8512	13.734

M1: $M_t = \beta_0 + \beta_1 t + w_t$

M2: $M_t = \beta_0 + \beta_1 t + \beta_2 (L_t - L.) + w_t$

M3: $M_t = \beta_0 + \beta_1 t + \beta_2 (L_t - L.) + \beta_3 (L_t - L.)^2 + w_t$

Where L_t = Population & $L.$ = Mean population

27

Concerns

- **Model 1, 2** – Q-stats indicate that even transformed red meat is not white noise; a better model is needed.
  - Alternative modeling techniques?
  - Alternative transformations of total red meat production?
- **Model 3** – Personal Income had collinearity issues; Population was not significant in the model.
  - Alternative covariates that may be related to meat production
  - Environmental variables, retail value of red meat, etc.

28

Discussion & Conclusion

- Despite our best efforts and fitting over 50 models, we could not achieve whiteness (even after considering more complex models) in ARIMA or SARIMA models
- Our covariates were not significant predictors of red meat production, which made investigating regression with autocorrelated errors difficult.
- Even after models were fit to account for autocorrelated variables, population did not become significant.

29

Questions?

30

Appendix

1. *USDA ERS - Livestock & Meat Domestic Data*. https://www.ers.usda.gov/data-products/livestock-meat-domestic-data/. Accessed 13 June 2021.
2. Shumway, Robert, and David Stoffer. *Time Series: A Data Analysis Approach Using R (Chapman & Hall/CRC Texts in Statistical Science)*. 1st ed., Chapman and Hall/CRC, 2019.
3. Thornton, Philip K. "Livestock production: recent trends, future prospects." *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* vol. 365,1554 (2010): 2853-67. doi:10.1098/rstb.2010.0134
4. US Bureau of Economic Analysis, Personal Income [PI], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/PI, July 4, 2021.
5. US Bureau of Economic Analysis, Population [POPTHM], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/POPTHM, July 4, 2021

31