Price relations of food commodities in regional markets of Nepal

Samita Paudel

6/28/2019

–>

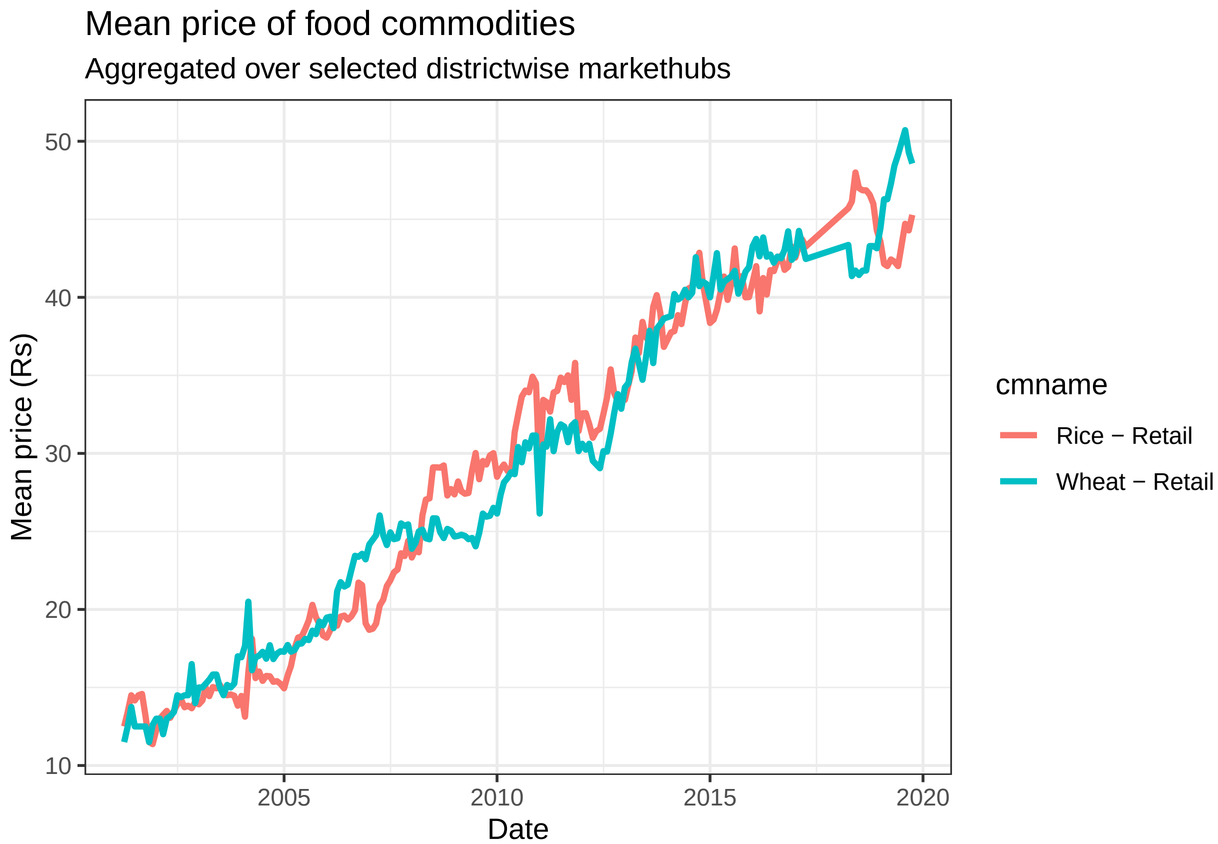
1 Retail price of rice and wheat in major districtwise Nepalese market hubs

The price series of districtwise market hubs available for study is mostly imbalanced and irregular and contains data for following 21 districts.

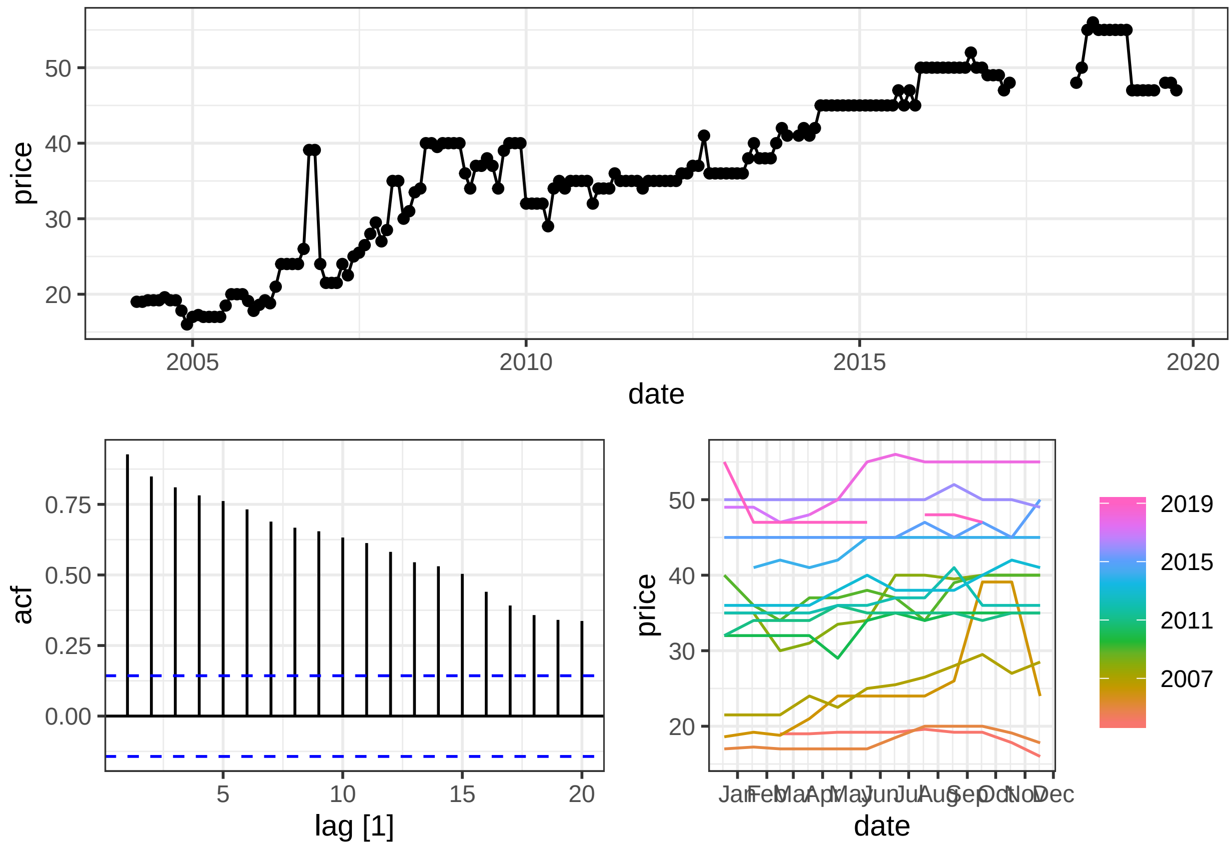
Achham, Banke, Bhojpur, Chitwan, Dhankuta, Dhanusha, Doti, Illam, Jhapa, Jumla, Kailali, Kaski, Kathmandu, Morang, Nuwakot, Palpa, Parsa, Ramechap, Rolpa, Rupandehi, Surkhet.

1.1 Aggregate series summary

Joint time series plot of rice and wheat retail prices aggregated over selected districtwise markets.

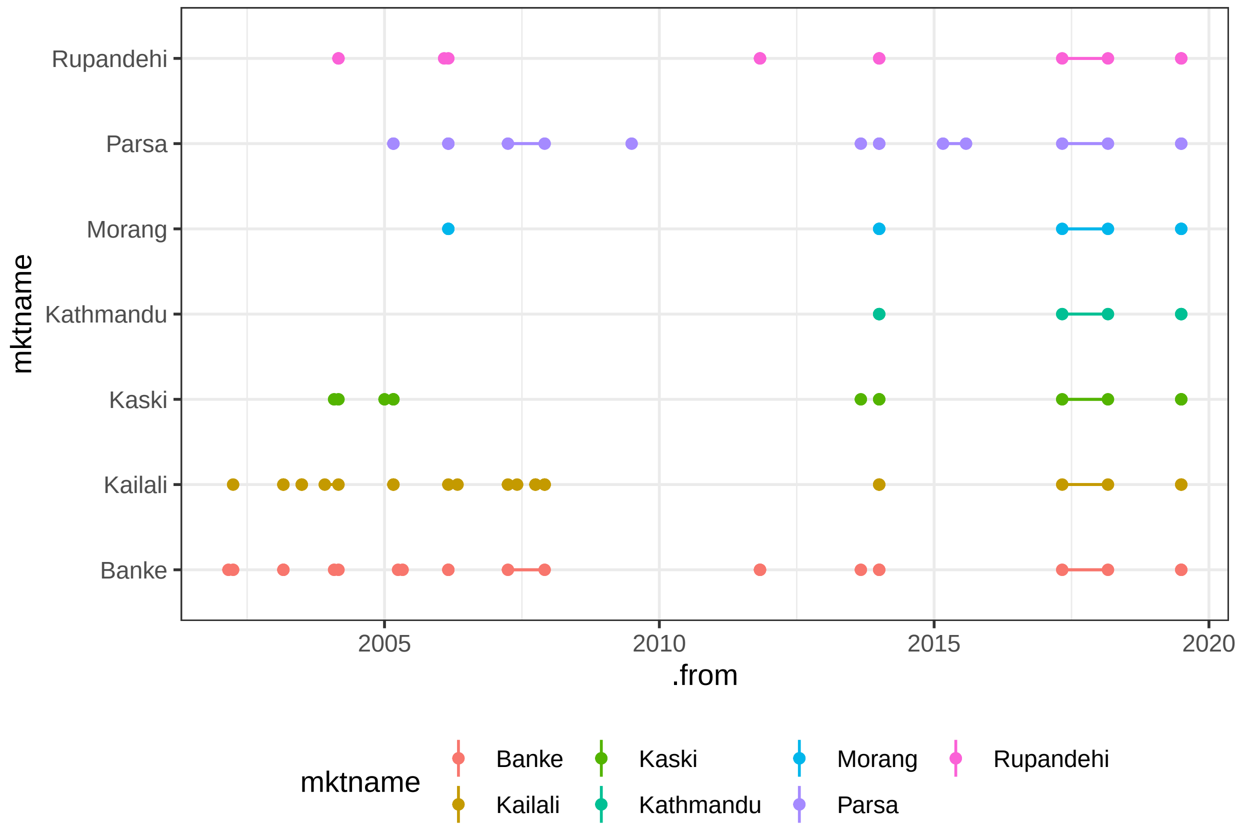


Time series plot of retail price of rice in Kathmandu market. Series has some time gaps at random periods (shown on line plot on the lower right). Similarly, autocorrelation of series for various lag, with first order difference, is presented in the lower left.

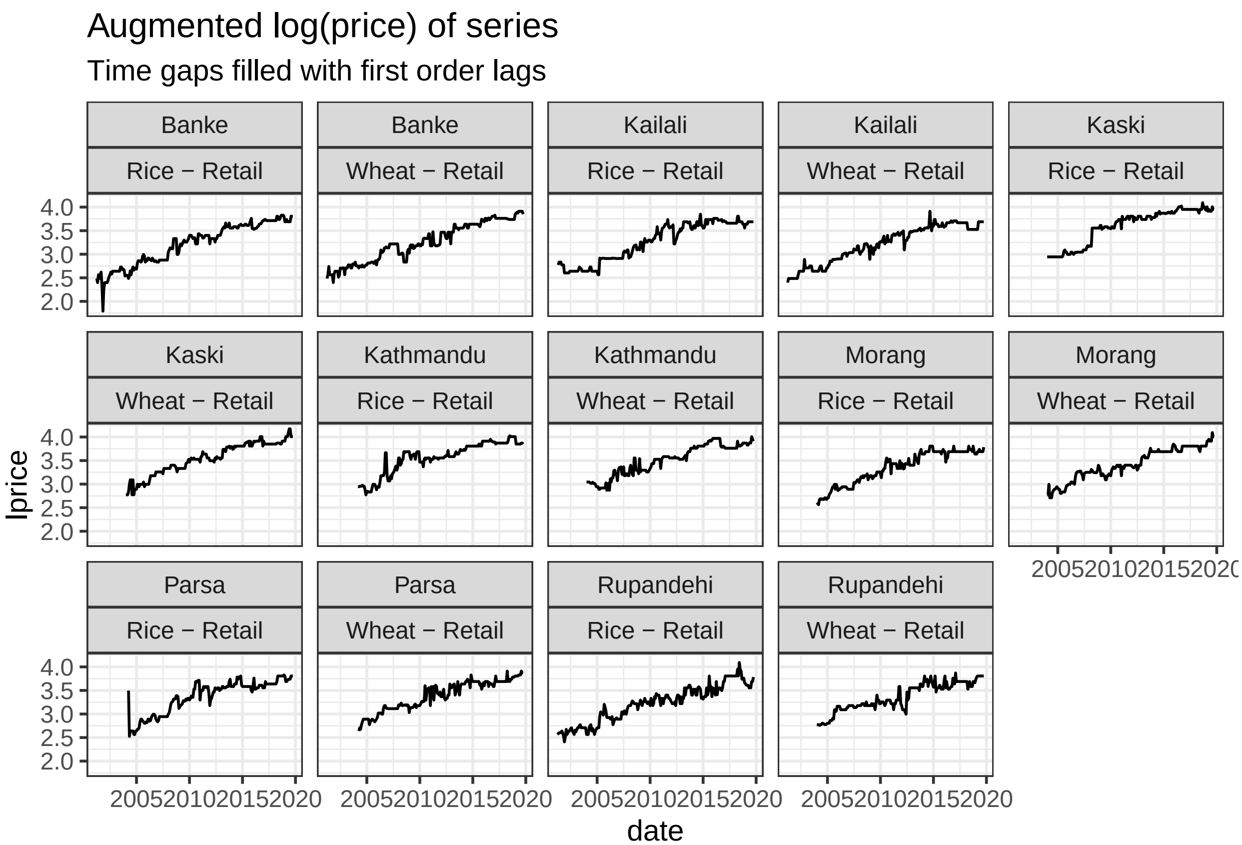


A possible measure to removing non-stationary trend in the series is by differencing (with `diff`) . However, before progressing we confirm that justifiable lag operations can infact render the series free of trends. For this at two fundamentally different unit tests are performed – Augmented Dickey-Fueller test and KPSS test.

While the series needs detrending in order to perform regression, we also have to consider the time gaps in the available dataset.



All district market series presented after time gap filling. First order lag is used to fill the missing entries in the series. Missing values for each series are imputed independent of other series.



2 Unit root testing

2.1 Unit root (ADF and KPSS) test of retail price

The ADF, available in the function `adf.test()` (in the package `tseries`) implements the t-test of $H_0 : \gamma = 0$ in the regression, below.

$$\Delta Y_t = \beta_1 + \beta_2 t + \gamma Y_{t-1} + \sum_{i=1}^{m} \delta_i \Delta Y_{t-i} + \varepsilon_t \tag{1}$$

The null is therefore that x has a unit root. If only x has a non-unit root, then the x is stationary (rejection of null hypothesis).

The ADF test was parametrized with the alternative hypothesis of stationarity. This extends to following assumption in the model parameters;

$$-2 \leq \gamma \leq 0 \text{ or } (-1 < 1 + \phi < 1)$$

`k` in the function refers to the number of δ lags, i.e., $1, 2, 3, ..., m$ in the model equation.

Table 1: Unit root test of district market retail log(price) data of Rice

mktname	adf pvalue	adf tstatistic	adf stationary	kpss pvalue	kpss tstatistic	kpss stationary
Banke	0.22	-2.86	FALSE	0.01	3.70	FALSE
Kailali	0.52	-2.13	FALSE	0.01	3.72	FALSE
Kaski	0.86	-1.33	FALSE	0.01	3.20	FALSE
Kathmandu	0.45	-2.30	FALSE	0.01	3.00	FALSE
Morang	0.44	-2.32	FALSE	0.01	3.42	FALSE
Parsa	0.22	-2.86	FALSE	0.01	2.64	FALSE
Rupandehi	0.01	-4.51	TRUE	0.01	3.91	FALSE

Table 2: Unit root test of district market retail log(price) data of Wheat

mktname	adf pvalue	adf tstatistic	adf stationary	kpss pvalue	kpss tstatistic	kpss stationary
Banke	0.01	-4.09	TRUE	0.01	3.86	FALSE
Kailali	0.47	-2.26	FALSE	0.01	3.81	FALSE
Kaski	0.05	-3.41	FALSE	0.01	3.40	FALSE
Kathmandu	0.34	-2.56	FALSE	0.01	3.43	FALSE
Morang	0.03	-3.62	TRUE	0.01	3.40	FALSE
Parsa	0.04	-3.55	TRUE	0.01	3.05	FALSE
Rupandehi	0.02	-3.72	TRUE	0.01	3.16	FALSE

The number of lags `k` defaults to `trunc((length(x)-1)^(1/3))`, where `x` is the series being tested. The default value of `k` corresponds to the suggested upper bound on the rate at which the number of lags, `k`, should be made to grow with the sample size for the general ARMA(p,q) setup `citation(package = "tseries")`.

For a Dickey-Fueller test, so only up to AR(1) time dependency in our stationary process, we set `k = 0`. Hence we have no δs (lags) in our test.

The DF model can be written as:

$$Y_t = \beta_1 + \beta_2 t + \phi Y_{t-1} + \varepsilon_t$$

It can be re-written so we can do a linear regression of ΔY_t against t and Y_{t-1} and test if ϕ is different from 0. If only, ϕ is not zero and assumption above $(-1 < 1 + \phi < 1)$ holds, the process is stationary. If ϕ is straight up 0, then we have a random walk process – all white noise.

$$\Delta Y_t = \beta_1 + \beta_2 t + \gamma Y_{t-1} + \varepsilon_t$$

Alternative to above discussed tests, the Phillips-Perron test with its nonparametric correction for autocorrelation (essentially employing a HAC estimate of the long-run variance in a Dickey-Fuller-type test instead of parametric decorrelation) may be used. It is available in the function `pp.test()`.

2.2 Unit root test based lag order differencing determination

An alternative to decomposition for removing trends is differencing (Woodward, Gray, and Elliott 2017). We define the difference operator as,

$$\nabla x_t = x_t - x_{t-1}, \tag{2}$$

and, more generally, for order d

$$\nabla^d x_t = (1 - \mathbf{B})^d x_t, \tag{3}$$

Where $\mathbf{B}$ is the backshift operator (i.e., $\mathbf{B}^k x_t = x_{t-k}$ for $k \geq 1$).

Applying the difference to a random walk, the most simple and widely used time series model, will yield a time series of Gaussian white noise errors $\{w_t\}$:

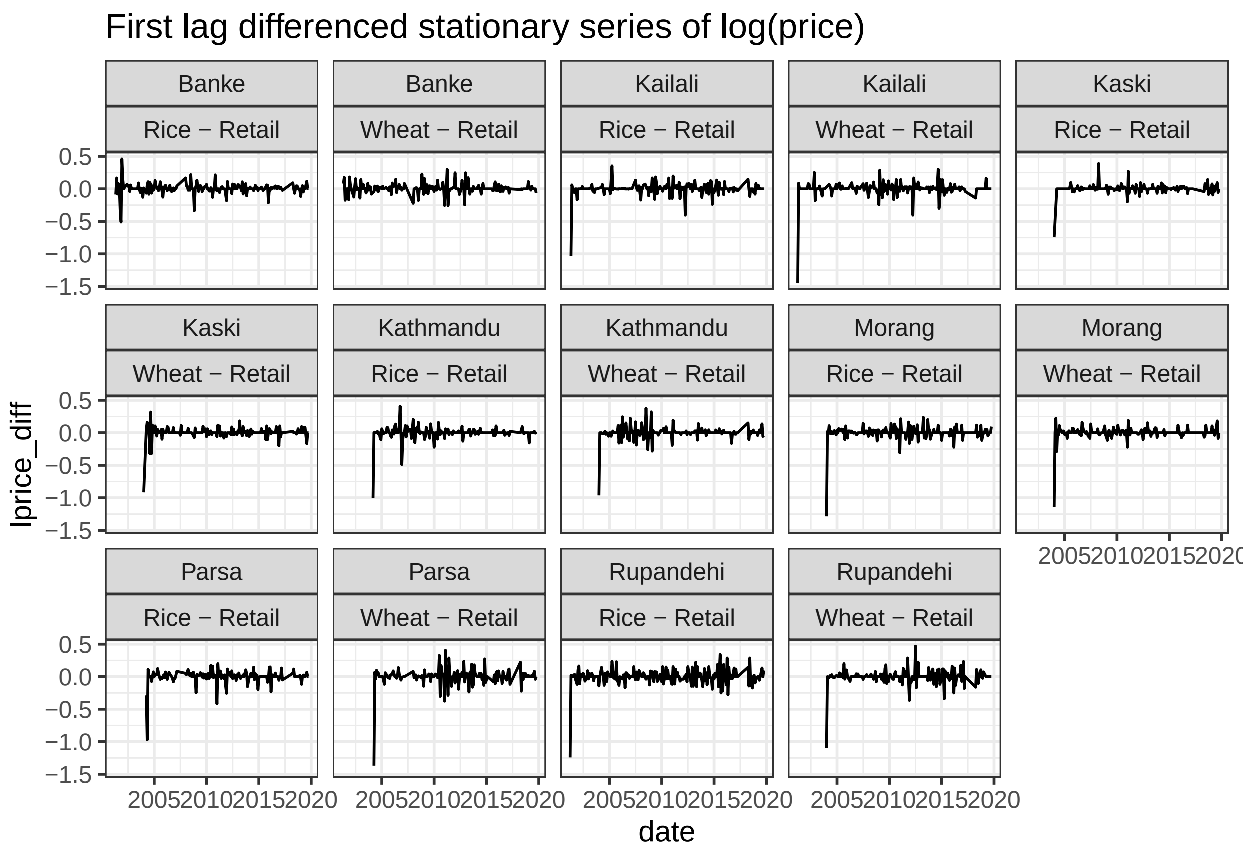
$$\begin{aligned} \nabla(x_t = x_{t-1} + w_t) \\ x_t - x_{t-1} = x_{t-1} - x_{t-1} + w_t \\ x_t - x_{t-1} = w_t \end{aligned} \tag{4}$$

Differencing is required for *all* series to make them stationary, as inferred by the `ndiffs` function which employed popular unit root tests. A detailed presentation of the test routines is given below. We describe in detail the ADF test, while only tabluation of summary statistics of `kpss` test is made herein.

2.3 Unit root tests of first order differenced series

All major districtwise market series series are non-stationary, meaning that they have a trend associated with time.

We test the logged prices of the series after first order differencing. Here we perform a more conservative Dickey-Fueller, instead of Augmented DF, test.



The first order differencing renders all series stationary.

2.4 ARIMA modeling log(retail) price

2.4.1 AR model

A simple way to model dependence between consecutive observations is by specification of model such as one in Equation 5.

$$y_t = \mu_0 + \phi_1 y_{t-1} + \varepsilon_t \tag{5}$$

Where, ε_t is the white noise having constant mean and variance and is independently distributed, i.e., has no autocorrelation.

Model representation of equation 5 is called a first-order autoregressive process ($AR(1)$), as it has one lag of y_t.

A p^{th} order autoregressive model – $AR(p)$ can be defined as in Equation 6.

$$y_t = \mu_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + ... + \phi_p y_{t-p} + \varepsilon_t \tag{6}$$

Autoregressive process of order (p) can also be represented using lag operator as in Equation 7.

$$y_t = \mu_0 + \sum_{i=1}^{p} \phi_i L^i y_t + \varepsilon_t \tag{7}$$

Where,

$$Ly_t = y_{t-1},\ L^2 y_t = y_{t-2},\ L^i y_t = y_{t-i}$$

For an $AR(1)$ model (Equation 5), following states may exist:

- When $\phi_1 = 0$, y_t is equivalent to white noise;
- When $\phi_1 = 1$ and $\mu_0 = 0$, y_t is equivalent to random walk;
- When $\phi_1 = 1$ and $c \neq 0$, y_t is equivalent to a random walk with drift;
- When $\phi_1 < 0$, y_t tends to oscillate between positive and negative value.

2.4.2 MA model

An MA(q) process is

$$y_t = c + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + ... + \theta_q \varepsilon_{t-q} \tag{8}$$

Following assumptions may hold for a MA(q) model

- For an MA(1) model: $-1 < \theta_1 < 1$
- For an MA(2) model: $-1 < \theta_2 < 1$, $\theta_2 + \theta_1 > -1$, $\theta_1 - \theta_2 < 1$

A process x_t is said to be ARIMA(p, d, q) if

$$\nabla^d x_t = (1 - B)^d x_t \tag{9}$$

Table 3: ARIMA model summary for multiple log(price) series of major districtwise market hubs of Rice - Retail

mktname	sigma2	log_lik	AIC	AICc	BIC
Kathmandu	0.004	243.522	-479.045	-478.825	-466.120
Parsa	0.010	167.945	-331.890	-331.824	-325.438
Morang	0.004	257.864	-509.728	-509.598	-500.003
Kailali	0.004	290.087	-568.173	-567.781	-547.784
Banke	0.005	265.152	-520.305	-520.027	-503.291
Kaski	0.002	304.876	-603.752	-603.622	-594.027
Rupandehi	0.007	234.554	-463.107	-462.997	-452.899

Table 4: ARIMA model summary for multiple log(price) series of major districtwise market hubs of Wheat - Retail

mktname	sigma2	log_lik	AIC	AICc	BIC
Kathmandu	0.005	239.962	-473.923	-473.793	-464.198
Parsa	0.006	212.092	-418.184	-418.052	-408.507
Morang	0.003	282.470	-554.939	-554.611	-538.730
Kailali	0.004	293.371	-578.742	-578.558	-565.131
Banke	0.004	291.350	-568.700	-568.176	-544.881
Kaski	0.003	274.450	-540.900	-540.682	-527.933
Rupandehi	0.006	212.393	-414.786	-414.458	-398.578

if ARMA(p, q). In general, we write the total model as:

$$\phi(B)(1-B)^d x_t = \theta(B) w_t \tag{10}$$

If $E(\nabla^d x_t) = \mu$, we write the model as:

$$\phi(B)(1-B)^d x_t = \delta + \theta(B) w_t \tag{11}$$

Where $\delta = \mu(1 - \phi_1 - ... - \phi_p)$.

3 VAR modeling

VAR is a system regression model, i.e., there are more than one dependent variable. The regression is defined by a set of linear dynamic equations where each variable is specified as a function of an equal number of lags of itself and all other variables in the system. Any

Table 5: Model coefficients of ARIMA model for multiple log(price) series of major district-wise market hubs of Rice - Retail

mktname	term	estimate	std.error	statistic	p.value
Kathmandu	ma1	0.025	0.072	0.350	0.727
Kathmandu	ma2	-0.353	0.074	-4.789	0.000
Kathmandu	constant	0.005	0.003	1.512	0.132
Parsa	ar1	-0.198	0.104	-1.912	0.057
Morang	ma1	-0.403	0.079	-5.096	0.000
Morang	constant	0.006	0.003	2.311	0.022
Kailali	ar1	1.387	NaN	NaN	NaN
Kailali	ar2	-0.485	NaN	NaN	NaN
Kailali	ma1	-1.579	NaN	NaN	NaN
Kailali	ma2	0.617	NaN	NaN	NaN
Kailali	constant	0.000	0.000	2.511	0.013
Banke	ar1	0.702	0.065	10.752	0.000
Banke	ma1	-0.947	0.030	-32.040	0.000
Banke	sar1	-0.013	0.093	-0.141	0.888
Banke	constant	0.002	0.000	6.728	0.000
Kaski	ma1	-0.316	0.070	-4.542	0.000
Kaski	constant	0.005	0.002	2.233	0.027
Rupandehi	ma1	-0.359	0.063	-5.701	0.000
Rupandehi	sma1	0.100	0.064	1.557	0.121

Table 6: Model coefficients of ARIMA model for multiple log(price) series of major district-wise market hubs of Wheat - Retail

mktname	term	estimate	std.error	statistic	p.value
Kathmandu	ma1	-0.499	0.066	-7.600	0.000
Kathmandu	constant	0.005	0.002	1.901	0.059
Parsa	ma1	-0.618	0.061	-10.188	0.000
Parsa	constant	0.006	0.002	2.906	0.004
Morang	ma1	-0.235	0.072	-3.249	0.001
Morang	ma2	-0.107	0.077	-1.384	0.168
Morang	sar1	0.118	0.083	1.418	0.158
Morang	constant	0.006	0.003	2.249	0.026
Kailali	ar1	0.276	0.148	1.865	0.063
Kailali	ma1	-0.700	0.116	-6.017	0.000
Kailali	constant	0.004	0.001	3.111	0.002
Banke	ma1	-0.310	0.066	-4.678	0.000
Banke	ma2	-0.213	0.074	-2.872	0.004
Banke	sar1	-0.746	0.315	-2.366	0.019
Banke	sar2	-0.183	0.074	-2.481	0.014
Banke	sma1	0.647	0.323	2.002	0.047
Banke	constant	0.012	0.003	3.299	0.001
Kaski	ar1	-0.412	0.066	-6.196	0.000
Kaski	sar1	0.093	0.097	0.954	0.341
Kaski	constant	0.008	0.004	2.066	0.040
Rupandehi	ar1	-0.715	0.163	-4.393	0.000
Rupandehi	ma1	0.403	0.158	2.554	0.011
Rupandehi	ma2	-0.381	0.070	-5.480	0.000
Rupandehi	sma1	0.215	0.075	2.854	0.005

additional variable, adds to the modeling complexity by increasing an extra equation to be estimated.

The vector autoregression (VAR) model extends the idea of univariate autoregression to k time series regressions, where the lagged values of *all* k series appear as regressors. Put differently, in a VAR model we regress a *vector* of time series variables on lagged vectors of these variables. As for $AR(p)$ models, the lag order is denoted by p so the $VAR(p)$ model of two variables X_t and Y_t ($k = 2$) is given by a vector of equations (Equation 12).

$$\begin{aligned} Y_t &= \beta_{10} + \beta_{11} Y_{t-1} + \cdots + \beta_{1p} Y_{t-p} + \gamma_{11} X_{t-1} + \cdots + \gamma_{1p} X_{t-p} + u_{1t}, \\ X_t &= \beta_{20} + \beta_{21} Y_{t-1} + \cdots + \beta_{2p} Y_{t-p} + \gamma_{21} X_{t-1} + \cdots + \gamma_{2p} X_{t-p} + u_{2t}. \end{aligned} \tag{12}$$

The βs and γs can be estimated using OLS on each equation.

Simplifying this to a bivariate $VAR(1)$, we can write the model in matrix form as:

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \mu_t \tag{13}$$

Where,

- Y_t, Y_{t-1} and μ_t are (2 x 1) column vectors
- β_0 is a (2 x 1) column vector
- β_1 is a (2 x 2) matrix

also,

$$Y_t = \begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix}, \ Y_{t-1} = \begin{pmatrix} y_{1t-1} \\ y_{2t-1} \end{pmatrix}$$

$$\mu_t = \begin{pmatrix} \mu_{1t} \\ \mu_{2t} \end{pmatrix}, \beta_0 = \begin{pmatrix} \beta_{10} \\ \beta_{20} \end{pmatrix}, \beta_1 = \begin{pmatrix} \beta_{11} & \alpha_{11} \\ \alpha_{21} & \beta_{21} \end{pmatrix}$$

It is straightforward to estimate VAR models in `R`. A feasible approach is to simply use `lm()` for estimation of the individual equations. Furthermore, the `vars` package provides standard tools for estimation, diagnostic testing and prediction using this type of models.

Only when the assumptions presented below hold, the OLS estimators of the VAR coefficients are consistent and jointly normal in large samples so that the usual inferential methods such as confidence intervals and t-statistics can be used (Metcalfe and Cowpertwait 2009).

Two series $w_{x,t}$ and $w_{y,t}$ are bivariate white noise if they are stationary and their cross-covariances $\gamma_{xy}(k) = Cov(w_{x,t}, w_{y,t+k})$ satisfies

$$\gamma_{xx}(k) = \gamma_{yy}(k) = \gamma_{xy}(k) = 0 \text{ for all } k \neq 0$$

The parameters of a var(p) model can be estimated using the `ar` function in `R`, which selects a best-fitting order p based on the smallest information criterion values.

The structure of VARs also allows to jointly test restrictions across multiple equations. For instance, it may be of interest to test whether the coefficients on all regressors of the lag p are zero. This corresponds to testing the null that the lag order $p-1$ is correct. Large sample joint normality of the coefficient estimates is convenient because it implies that we may simply use an F-test for this testing problem. The explicit formula for such a test statistic is rather complicated but fortunately such computations are easily done using the `ttcode("R")` functions we work with in this chapter. Just as in the case of a single equation, for a multiple equation model we choose the specification which has the smallest $BIC(p)$, where

$$BIC(p) = \log\left[\det(\widehat{\Sigma}_u)\right] + k(kp+1)\frac{\log(T)}{T}.$$

with $\widehat{\Sigma}_u$ denoting the estimate of the $k \times k$ covariance matrix of the VAR errors and $\det(\cdot)$ denotes the determinant.

As for univariate distributed lag models, one should think carefully about variables to include in a VAR, as adding unrelated variables reduces the forecast accuracy by increasing the estimation error. This is particularly important because the number of parameters to be estimated grows qudratically to the number of variables modeled by the VAR.

Table 7: Model performance indicators of VAR(AR(1)) model for selected districtwise market hub series in Rice and Wheat.

cmname-mktname	sigma2	log_lik	AIC	AICc	BIC
Rice - Retail / Banke	0.005987452	254.108	-502.216	-502.106	-492.008
Rice - Retail / Kailali	0.004562652	282.494	-560.989	-560.934	-554.192
Rice - Retail / Kaski	0.002558304	296.841	-587.682	-587.552	-577.956
Rice - Retail / Kathmandu	0.004799157	234.890	-463.780	-463.649	-454.086
Rice - Retail / Morang	0.004287195	248.052	-490.103	-489.973	-480.378
Rice - Retail / Parsa	0.009750488	167.714	-329.427	-329.295	-319.750
Rice - Retail / Rupandehi	0.007998555	221.963	-437.926	-437.816	-427.718
Wheat - Retail / Banke	0.004799646	278.653	-551.306	-551.196	-541.098
Wheat - Retail / Kailali	0.004939438	275.466	-544.932	-544.822	-534.724
Wheat - Retail / Kaski	0.003879621	257.492	-508.983	-508.854	-499.258
Wheat - Retail / Kathmandu	0.005504843	224.427	-442.854	-442.724	-433.129
Wheat - Retail / Morang	0.003155432	277.016	-548.033	-547.903	-538.308
Wheat - Retail / Parsa	0.007953767	186.655	-367.310	-367.178	-357.633
Wheat - Retail / Rupandehi	0.007253182	198.363	-390.725	-390.596	-381.000

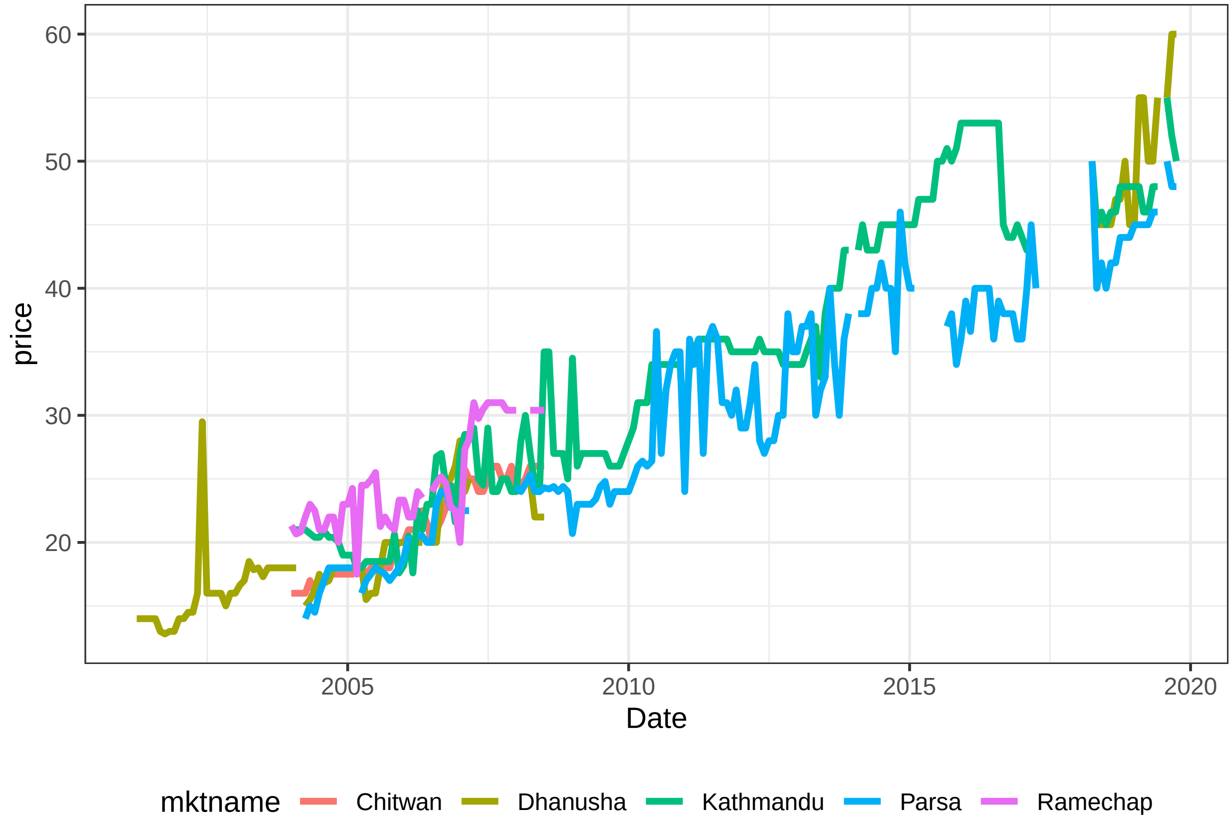
Table 8: Model coefficients of VAR model for multiple log(price) series of major districtwise market hubs of Rice - Retail

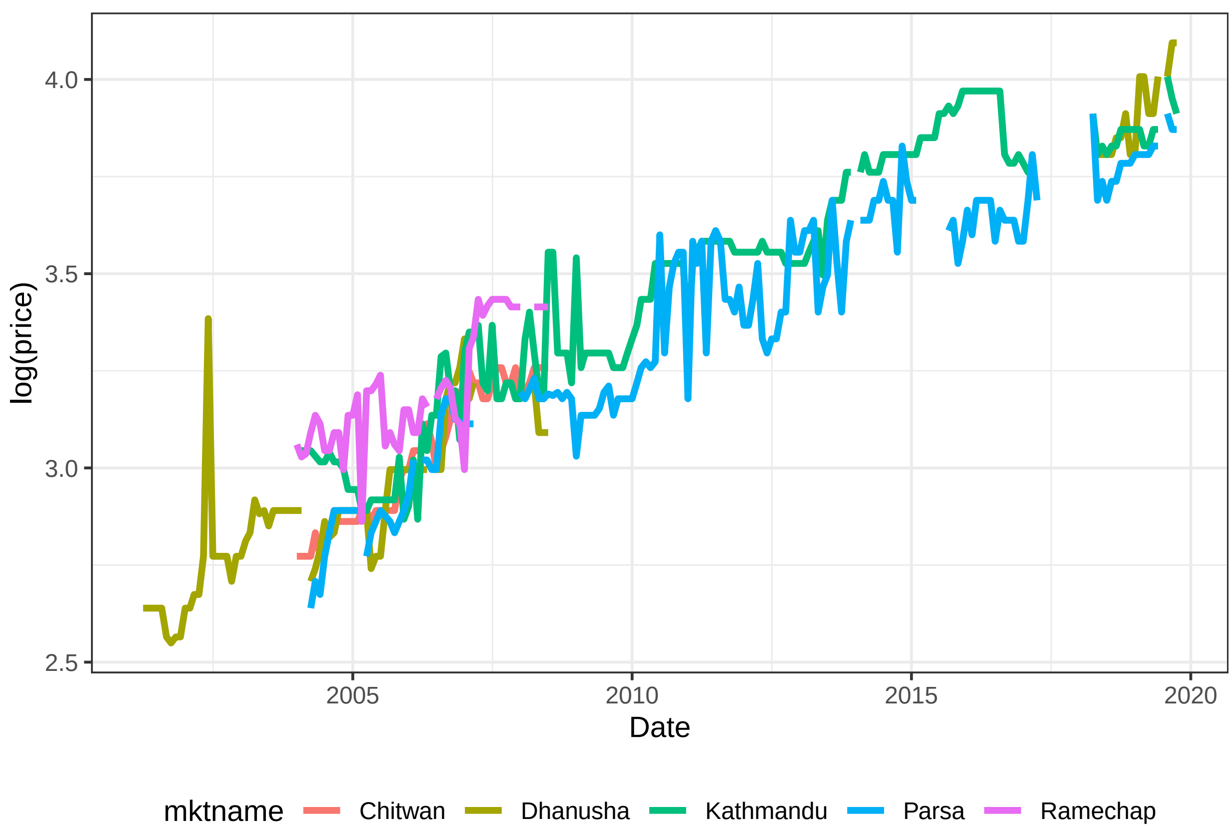
mktname	term	estimate	std.error	statistic	p.value
Banke	lag(lprice,1)	0.982	0.012	82.933	0.000
Banke	constant	0.062	0.038	1.626	0.105
Kailali	lag(lprice,1)	1.001	0.001	723.506	0.000
Kaski	lag(lprice,1)	0.985	0.010	99.189	0.000
Kaski	constant	0.059	0.036	1.654	0.100
Kathmandu	lag(lprice,1)	0.971	0.015	63.567	0.000
Kathmandu	constant	0.109	0.055	1.999	0.047
Morang	lag(lprice,1)	0.973	0.014	71.100	0.000
Morang	constant	0.098	0.046	2.121	0.035
Parsa	lag(lprice,1)	0.963	0.021	46.349	0.000
Parsa	constant	0.125	0.070	1.774	0.078
Rupandehi	lag(lprice,1)	0.972	0.016	62.573	0.000
Rupandehi	constant	0.097	0.050	1.927	0.055

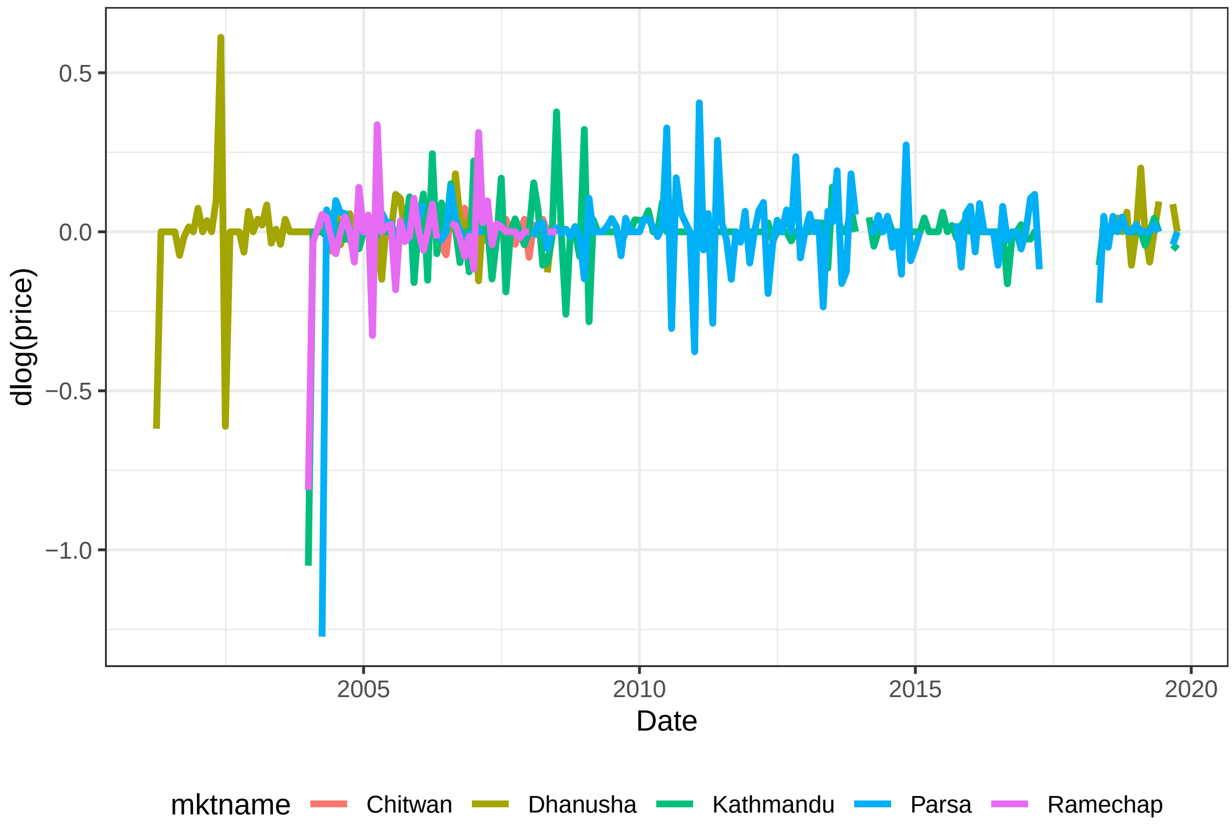
Table 9: Model coefficients of VAR model for multiple log(price) series of major districtwise market hubs of Wheat - Retail

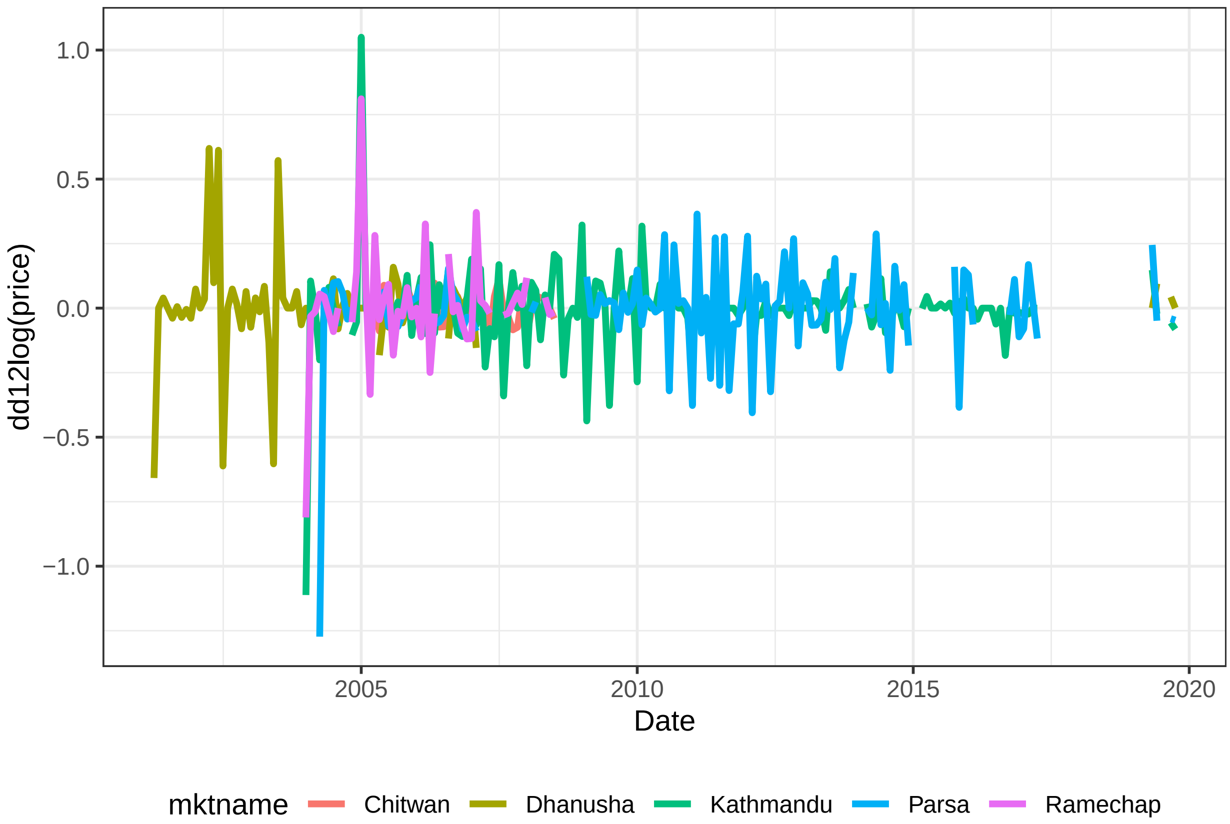
mktname	term	estimate	std.error	statistic	p.value
Banke	lag(lprice,1)	0.982	0.011	86.538	0.000
Banke	constant	0.063	0.037	1.698	0.091
Kailali	lag(lprice,1)	0.977	0.012	80.478	0.000
Kailali	constant	0.080	0.039	2.033	0.043
Kaski	lag(lprice,1)	0.974	0.013	72.988	0.000
Kaski	constant	0.097	0.047	2.046	0.042
Kathmandu	lag(lprice,1)	0.972	0.017	57.967	0.000
Kathmandu	constant	0.103	0.059	1.741	0.083
Morang	lag(lprice,1)	0.983	0.012	79.565	0.000
Morang	constant	0.063	0.043	1.492	0.137
Parsa	lag(lprice,1)	0.949	0.021	45.575	0.000
Parsa	constant	0.178	0.071	2.514	0.013
Rupandehi	lag(lprice,1)	0.955	0.021	46.210	0.000
Rupandehi	constant	0.157	0.070	2.259	0.025

3.1 Differenced series






0.5
0.0
−0.5
−1.0
dlog(price)
2005
2010
2015
2020
Date
mktname
Chitwan
Dhanusha
Kathmandu
Parsa
Ramechap



4 Cointegration

4.1 Residual based

Since the food commodities are spatially linked, more of so because they occupy the same domestic market, it is obvious that factor affecting price of one inevitably affects other, especially that of same crop in a nearby market. Having evidence for nonstationarity, it is of interest to test for a common nonstationary component by means of a cointegration test (Non-stationarity is more valid for development regionwise price series).

A two step method proposed by Hylleberg et al. (1990) can be used to test for cointegration.

The procedure simply regressess one series on the other and performs a unit root test on the residuals. This test is often named after Phillips, Ouliaris, and others (1990). Specifically, `po.test()` performs a Phillips-Perron test using an auxiliary regression without a constant and linear trend and the Newey-West estimator for the required long-run variance.

The test computes the Phillips-Ouliaris test for the null hypothesis that series is not cointegrated (Trapletti and Hornik 2019).

We check the rice retail price series for all combination major districtwise markets.

Table 10: Phillips-Ouliaris cointegration test for Rice log(price) series of selected district markethubs

combination	p_value	statistic
Morang-Parsa	0.010	-96.515
Morang-Kathmandu	0.010	-34.869
Morang-Kaski	0.010	-37.126
Morang-Rupandehi	0.010	-42.435
Morang-Banke	0.010	-65.794
Morang-Kailali	0.010	-66.540
Parsa-Kathmandu	0.010	-59.925
Parsa-Kaski	0.010	-94.521
Parsa-Rupandehi	0.010	-63.256
Parsa-Banke	0.010	-91.480
Parsa-Kailali	0.010	-89.453
Kathmandu-Kaski	0.010	-35.060
Kathmandu-Rupandehi	0.010	-37.100
Kathmandu-Banke	0.010	-34.602
Kathmandu-Kailali	0.010	-34.579
Kaski-Rupandehi	0.013	-27.453
Kaski-Banke	0.010	-41.121
Kaski-Kailali	0.010	-52.289
Rupandehi-Banke	0.010	-89.619
Rupandehi-Kailali	0.010	-41.667
Banke-Kailali	0.010	-58.835

Table 11: Phillips-Ouliaris cointegration test for Wheat log(price) series of selected district markethubs

combination	p_value	statistic
Morang-Parsa	0.01	-79.626
Morang-Kathmandu	0.01	-53.276
Morang-Kaski	0.01	-71.222
Morang-Rupandehi	0.01	-79.788
Morang-Banke	0.01	-56.448
Morang-Kailali	0.01	-55.875
Parsa-Kathmandu	0.01	-80.952
Parsa-Kaski	0.01	-90.071
Parsa-Rupandehi	0.01	-78.571
Parsa-Banke	0.01	-57.703

Parsa-Kailali	0.01	-92.198
Kathmandu-Kaski	0.01	-82.232
Kathmandu-Rupandehi	0.01	-48.619
Kathmandu-Banke	0.01	-50.344
Kathmandu-Kailali	0.01	-73.156
Kaski-Rupandehi	0.01	-58.758
Kaski-Banke	0.01	-49.955
Kaski-Kailali	0.01	-87.482
Rupandehi-Banke	0.01	-57.133
Rupandehi-Kailali	0.01	-64.151
Banke-Kailali	0.01	-53.604

Note `po.test` does not handle missing values, so we fix them through imputation. It is implemented through `tidyr::fill(..., .direction = "down")`.

The test suggests that all series (Both that of wheat and rice) are cointegrated for selected pairwise combination of district markets.

The problem with this approach is that it treats both series in an asymmetric fashion, while the concept of cointegration demands that the treatment be symmetric.

The po.test() function is testing the cointegration with Phillip's Z_alpha test, which is the second residual-based test described by Phillips, Ouliaris, and others (1990). Because the po.test() will use the series at the first position to derive the residual used in the test, results would be determined by the series on the most left-hand side[1].

The Phillips-Ouliaris test implemented in the `ca.po()` function from the urca package is different. In the `ca.po()` function, there are two cointegration tests implemented, namely "Pu" and "Pz" tests. Although both the `ca.po()` function and the po.test() function are supposed to do the Phillips-Ouliaris test outcomes from both functions are completely different.

Similar to Phillip's Z_alpha test, the Pu test also is not invariant to the position of each series and therefore would give different outcomes based upon the series on the most left-hand side. On the contrary, the multivariate trace statistic of Pz test has its appeal in that the outcome won't change by the position of each series.

4.2 VAR based

The standard tests proceeding in a symmetric manner stem from Johansen's full-information maximum likelihood approach (Johansen 1991).

A general vector autoregressive model is similar to the AR(p) model except that each quantity is vector valued and matrices are used as the coefficients. The general form of the VAR(p) model, without drift, is given by:

[1] https://www.r-craft.org/r-news/phillips-ouliaris-test-for-cointegration/

$$\mathbf{y_t} = \quad + A_1\mathbf{y_{t-1}} + ... + A_j\mathbf{y_{t-j}} + \quad _{\mathbf{t}} \tag{14}$$

Where   is the vector-valued mean of the series, A_i are the coefficient matrices for each lag and $_{\mathbf{t}}$ is a multivariate Gaussian noise term with mean zero.

At this stage we can form a Vector Error Correction Model (VECM) by differencing the series (Equation 15).

$$\Delta\mathbf{y_t} = \quad + A\mathbf{y_{t-1}} + \Gamma_1\Delta\mathbf{y_{t-1}} + ... + \Gamma_j\Delta\mathbf{y_{t-j}} + \quad _{\mathbf{t}} \tag{15}$$

Where $\Delta\mathbf{y_t} = \mathbf{y_t} - \mathbf{y_{t-1}}$ is the differencing operator, A is the coefficient matrix for the first lag and Γ_i are the matrices for each differenced lag.

For a p^{th} -order cointegrated vector autoregressive (VAR) model, the error correction form is (omitting deterministic components; both no intercept or trend in either cointegrating equation or test var), we may rewrite the VAR in the form of Equation 16 (Johansen 1991).

$$\Delta y_t = \Pi y_{t-1} + \sum_{j=1}^{p-1}\Gamma_j\Delta y_{t-j} + \varepsilon_t \tag{16}$$

Where,

$$\Pi = \sum_{i=1}^{p} A_i - I; \Gamma = -\sum_{j=i+1}^{p} j$$

(Although, for simplicity sake, we assume absence of deterministic trends, there are five popular scenarios of including such trends in a cointegration test. All of these are described in (Johansen 1995).)

Granger's representation theorem asserts that if the coefficient matrix Π has reduced rank $r < k$, then there exist kxr matrices α and β each with rank k such that $\Pi = \alpha\beta'$ and $\beta' y_t$ is $I(0)$.

To achieve this an eigenvalue decomposition of A is carried out. The rank of the matrix A is given by r and the Johansen test sequentially tests whether this rank r is equal to zero, equal to one, through to $r = n - 1$, where n is the number of time series under test.

The null hypothesis of $r = 0$ means that there is no cointegration at all. A rank $r > 0$ implies a cointegrating relationship between two or possibly more time series.

The eigenvalue decomposition results in a set of eigenvectors. The components of the largest eigenvector admits the important property of forming the coefficients of a linear combination of time series to produce a stationary portfolio. Notice how this differs from the CADF test (often known as the Engle-Granger procedure) where it is necessary to ascertain the linear combination a priori via linear regression and ordinary least squares (OLS).

In summary, the test checks for the situation of no cointegration, which occurs when the matrix $A = 0$. So, starting with the base value of r (i.e., $r = 0$), if the test statistic is greater than critical values of at the 10%, 5% and 1% levels, this would imply that we are **able** to reject the null of no cointegration. For the case r<=1, we if the calculated test statistic is below the critical values of, we are **unable** to reject the null, and the number of cointegrating vectors is between 0 and 1. The relevant tests are available in the function `urca::ca.jo()`. The basic version considers the eigenvalues of the matrix Π in the preceding equation.

Here, we employ the trace statistic – the maximum eigenvalue, or “lambdamax” test is available as well – in an equation amended by a constant term (specified by ecdet = “const”), yielding:

Johansen cointegration test summary and time series plots for rice (district marketwise)

Table 12: Johansen cointegration test summary for Morang-Parsa

gamma	test_stat	10pct	5pct	1pct
r <= 1 |	8.047588	7.52	9.24	12.97
r = 0 |	29.352546	17.85	19.96	24.60

Table 13: Johansen cointegration test summary for Morang-Kathmandu

gamma	test_stat	10pct	5pct	1pct
r <= 1 |	7.080797	7.52	9.24	12.97
r = 0 |	26.213687	17.85	19.96	24.60

Table 14: Johansen cointegration test summary for Morang-Kaski

gamma	test_stat	10pct	5pct	1pct
r <= 1 |	10.47841	7.52	9.24	12.97
r = 0 |	26.58143	17.85	19.96	24.60

Table 15: Johansen cointegration test summary for Morang-Rupandehi

gamma	test_stat	10pct	5pct	1pct
r <= 1 |	9.529231	7.52	9.24	12.97
r = 0 |	28.702528	17.85	19.96	24.60

Table 16: Johansen cointegration test summary for Morang-Banke

gamma	test_stat	10pct	5pct	1pct
r <= 1 |	11.07168	7.52	9.24	12.97
r = 0 |	43.79799	17.85	19.96	24.60

Table 17: Johansen cointegration test summary for Morang-Kailali

gamma	test_stat	10pct	5pct	1pct
r <= 1 |	8.627746	7.52	9.24	12.97
r = 0 |	37.159947	17.85	19.96	24.60

Table 18: Johansen cointegration test summary for Parsa-Kathmandu

gamma	test_stat	10pct	5pct	1pct
r <= 1 |	7.117339	7.52	9.24	12.97
r = 0 |	28.121237	17.85	19.96	24.60

Table 19: Johansen cointegration test summary for Parsa-Kaski

gamma	test_stat	10pct	5pct	1pct
r <= 1 |	8.739686	7.52	9.24	12.97
r = 0 |	33.345326	17.85	19.96	24.60

Table 20: Johansen cointegration test summary for Parsa-Rupandehi

gamma	test_stat	10pct	5pct	1pct
r <= 1 |	5.241234	7.52	9.24	12.97
r = 0 |	21.237513	17.85	19.96	24.60

Table 21: Johansen cointegration test summary for Parsa-Banke

gamma	test_stat	10pct	5pct	1pct
r <= 1 |	8.905272	7.52	9.24	12.97
r = 0 |	35.297439	17.85	19.96	24.60

Table 22: Johansen cointegration test summary for Parsa-Kailali

gamma	test_stat	10pct	5pct	1pct
r <= 1 |	6.477453	7.52	9.24	12.97
r = 0 |	30.475019	17.85	19.96	24.60

Table 23: Johansen cointegration test summary for Kathmandu-Kaski

gamma	test_stat	10pct	5pct	1pct
r <= 1 |	6.896813	7.52	9.24	12.97
r = 0 |	30.474234	17.85	19.96	24.60

Table 24: Johansen cointegration test summary for Kathmandu-Rupandehi

gamma	test_stat	10pct	5pct	1pct
r <= 1 |	5.706426	7.52	9.24	12.97
r = 0 |	24.640127	17.85	19.96	24.60

Table 25: Johansen cointegration test summary for Kathmandu-Banke

gamma	test_stat	10pct	5pct	1pct
r <= 1 |	9.024811	7.52	9.24	12.97
r = 0 |	30.343046	17.85	19.96	24.60

Table 26: Johansen cointegration test summary for Kathmandu-Kailali

gamma	test_stat	10pct	5pct	1pct
r <= 1 |	6.166887	7.52	9.24	12.97
r = 0 |	27.654849	17.85	19.96	24.60

Table 27: Johansen cointegration test summary for Kaski-Rupandehi

gamma	test_stat	10pct	5pct	1pct
r <= 1 |	7.905289	7.52	9.24	12.97
r = 0 |	22.551933	17.85	19.96	24.60

Table 28: Johansen cointegration test summary for Kaski-Banke

gamma	test_stat	10pct	5pct	1pct
r <= 1 |	8.927975	7.52	9.24	12.97
r = 0 |	29.193633	17.85	19.96	24.60

Table 29: Johansen cointegration test summary for Kaski-Kailali

gamma	test_stat	10pct	5pct	1pct
r <= 1 |	7.744047	7.52	9.24	12.97
r = 0 |	31.015706	17.85	19.96	24.60

Table 30: Johansen cointegration test summary for Rupandehi-Banke

gamma	test_stat	10pct	5pct	1pct
r <= 1 |	4.441382	7.52	9.24	12.97
r = 0 |	46.391884	17.85	19.96	24.60

Table 31: Johansen cointegration test summary for Rupandehi-Kailali

gamma	test_stat	10pct	5pct	1pct
r <= 1 |	2.995481	7.52	9.24	12.97
r = 0 |	20.488277	17.85	19.96	24.60

Table 32: Johansen cointegration test summary for Banke-Kailali

gamma	test_stat	10pct	5pct	1pct
r <= 1 |	2.788889	7.52	9.24	12.97
r = 0 |	40.728755	17.85	19.96	24.60

Johansen cointegration test summary and time series plots for wheat (district marketwise)

Table 33: Johansen cointegration test summary for Morang-Parsa

gamma	test_stat	10pct	5pct	1pct
r <= 1 |	8.67185	7.52	9.24	12.97
r = 0 |	38.90466	17.85	19.96	24.60

Table 34: Johansen cointegration test summary for Morang-Kathmandu

gamma	test_stat	10pct	5pct	1pct
r <= 1 |	6.03380	7.52	9.24	12.97
r = 0 |	32.39395	17.85	19.96	24.60

Table 35: Johansen cointegration test summary for Morang-Kaski

gamma	test_stat	10pct	5pct	1pct
r <= 1 |	14.80835	7.52	9.24	12.97
r = 0 |	47.39019	17.85	19.96	24.60

Table 36: Johansen cointegration test summary for Morang-Rupandehi

gamma	test_stat	10pct	5pct	1pct
r <= 1 |	6.698168	7.52	9.24	12.97
r = 0 |	41.851350	17.85	19.96	24.60

Table 37: Johansen cointegration test summary for Morang-Banke

gamma	test_stat	10pct	5pct	1pct

r <= 1 |	6.791701	7.52	9.24	12.97
r = 0 |	29.981567	17.85	19.96	24.60

Table 38: Johansen cointegration test summary for Morang-Kailali

gamma	test_stat	10pct	5pct	1pct
r <= 1 |	10.09031	7.52	9.24	12.97
r = 0 |	34.84797	17.85	19.96	24.60

Table 39: Johansen cointegration test summary for Parsa-Kathmandu

gamma	test_stat	10pct	5pct	1pct
r <= 1 |	5.572257	7.52	9.24	12.97
r = 0 |	39.875312	17.85	19.96	24.60

Table 40: Johansen cointegration test summary for Parsa-Kaski

gamma	test_stat	10pct	5pct	1pct
r <= 1 |	9.047446	7.52	9.24	12.97
r = 0 |	43.549450	17.85	19.96	24.60

Table 41: Johansen cointegration test summary for Parsa-Rupandehi

gamma	test_stat	10pct	5pct	1pct
r <= 1 |	7.098776	7.52	9.24	12.97
r = 0 |	35.167352	17.85	19.96	24.60

Table 42: Johansen cointegration test summary for Parsa-Banke

gamma	test_stat	10pct	5pct	1pct
r <= 1 |	7.52789	7.52	9.24	12.97
r = 0 |	36.13913	17.85	19.96	24.60

Table 43: Johansen cointegration test summary for Parsa-Kailali

gamma	test_stat	10pct	5pct	1pct
r <= 1 |	8.944134	7.52	9.24	12.97
r = 0 |	48.858929	17.85	19.96	24.60

Table 44: Johansen cointegration test summary for Kathmandu-Kaski

gamma	test_stat	10pct	5pct	1pct
r <= 1 |	9.856107	7.52	9.24	12.97
r = 0 |	48.980703	17.85	19.96	24.60

Table 45: Johansen cointegration test summary for Kathmandu-Rupandehi

gamma	test_stat	10pct	5pct	1pct
r <= 1 |	4.913668	7.52	9.24	12.97
r = 0 |	27.111863	17.85	19.96	24.60

Table 46: Johansen cointegration test summary for Kathmandu-Banke

gamma	test_stat	10pct	5pct	1pct
r <= 1 |	5.079637	7.52	9.24	12.97
r = 0 |	27.171088	17.85	19.96	24.60

Table 47: Johansen cointegration test summary for Kathmandu-Kailali

gamma	test_stat	10pct	5pct	1pct
r <= 1 |	6.417786	7.52	9.24	12.97
r = 0 |	35.001568	17.85	19.96	24.60

Table 48: Johansen cointegration test summary for Kaski-Rupandehi

gamma	test_stat	10pct	5pct	1pct
r <= 1 |	10.27943	7.52	9.24	12.97
r = 0 |	33.21418	17.85	19.96	24.60

Table 49: Johansen cointegration test summary for Kaski-Banke

gamma	test_stat	10pct	5pct	1pct
r <= 1 |	8.023414	7.52	9.24	12.97
r = 0 |	32.713727	17.85	19.96	24.60

Table 50: Johansen cointegration test summary for Kaski-Kailali

gamma	test_stat	10pct	5pct	1pct
r <= 1 |	11.36236	7.52	9.24	12.97
r = 0 |	40.97431	17.85	19.96	24.60

Table 51: Johansen cointegration test summary for Rupandehi-Banke

gamma	test_stat	10pct	5pct	1pct
r <= 1 |	6.253642	7.52	9.24	12.97
r = 0 |	34.339634	17.85	19.96	24.60

Table 52: Johansen cointegration test summary for Rupandehi-Kailali

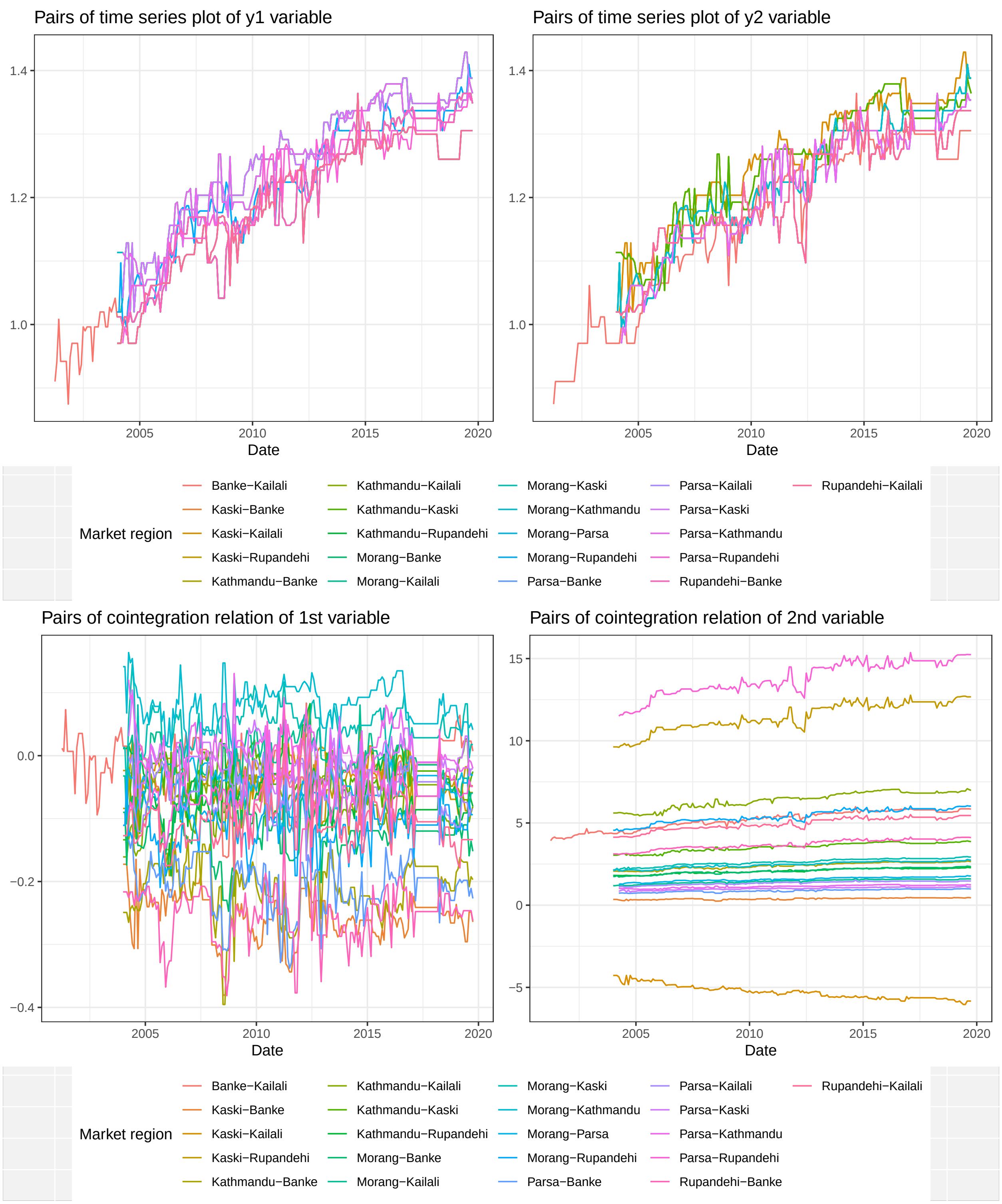
gamma	test_stat	10pct	5pct	1pct
r <= 1 |	7.004144	7.52	9.24	12.97
r = 0 |	31.782559	17.85	19.96	24.60

Table 53: Johansen cointegration test summary for Banke-Kailali

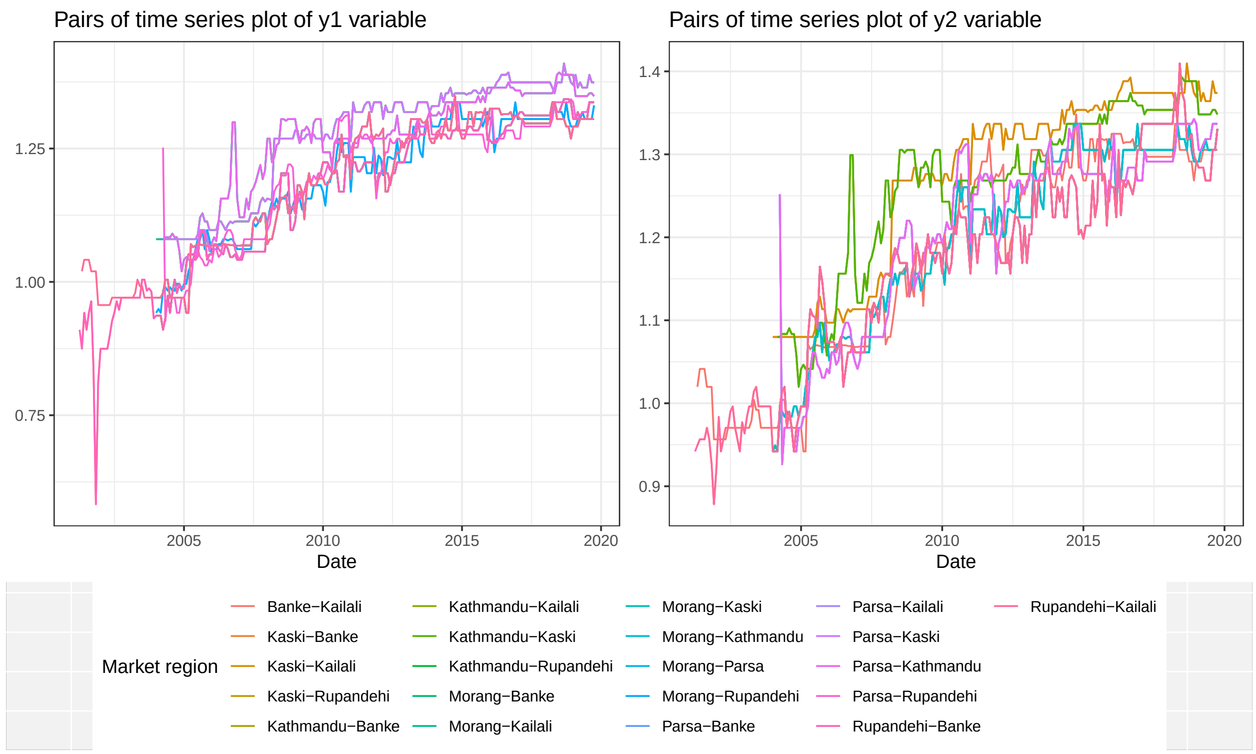
gamma	test_stat	10pct	5pct	1pct
r <= 1 |	7.875127	7.52	9.24	12.97

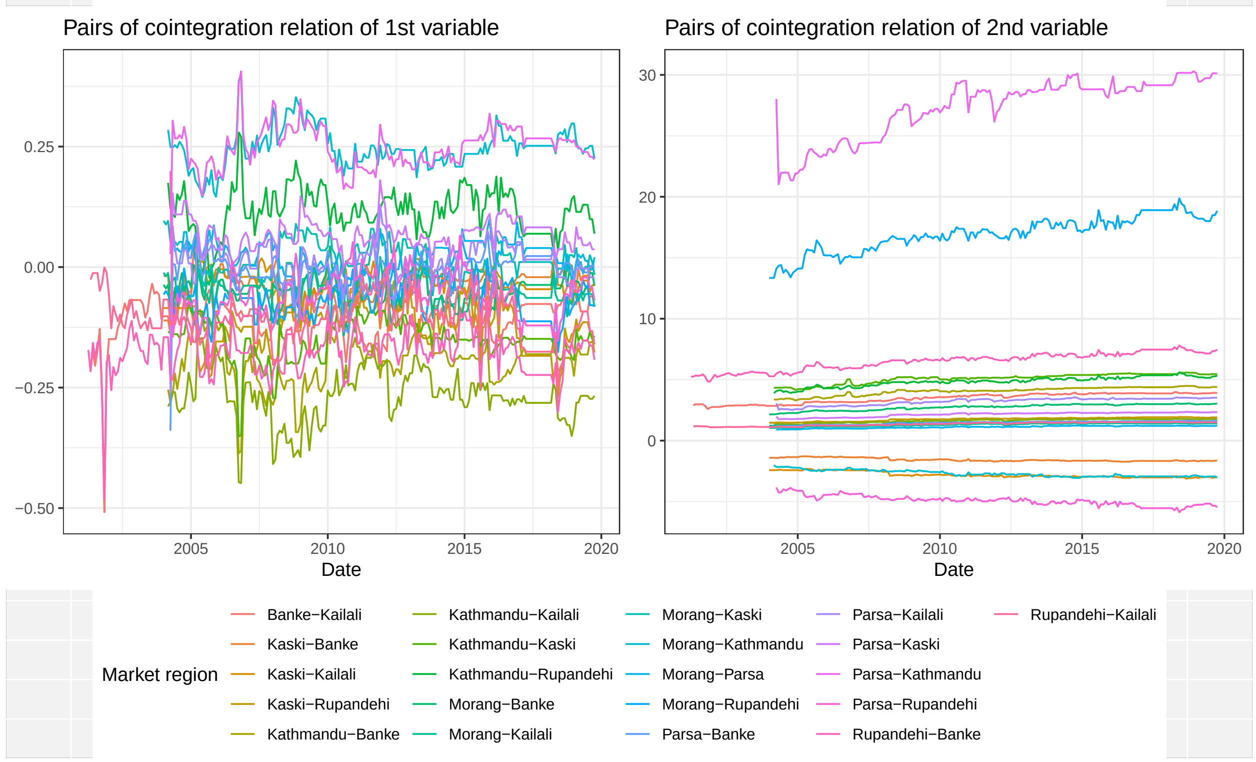
r = 0 |	32.936232	17.85	19.96	24.60

Wheat series and cointegration plots



Rice series and cointegration plots





Bibliography

Hylleberg, Svend, Robert F Engle, Clive WJ Granger, and Byung Sam Yoo. 1990. "Seasonal Integration and Cointegration." *Journal of Econometrics* 44 (1-2): 215–38.

Johansen, Søren. 1991. "Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models." *Econometrica: Journal of the Econometric Society*, 1551–80.

———. 1995. "Identifying Restrictions of Linear Equations with Applications to Simultaneous Equations and Cointegration." *Journal of Econometrics* 69 (1): 111–32.

Metcalfe, Andrew V, and Paul SP Cowpertwait. 2009. *Introductory Time Series with R*. Springer.

Phillips, Peter CB, Sam Ouliaris, and others. 1990. "Asymptotic Properties of Residual Based Tests for Cointegration." *Econometrica* 58 (1): 165–93.

Trapletti, Adrian, and Kurt Hornik. 2019. *Tseries: Time Series Analysis and Computational Finance*. https://CRAN.R-project.org/package=tseries.

Woodward, Wayne A, Henry L Gray, and Alan C Elliott. 2017. *Applied Time Series Analysis with R*. CRC press.