A multivariate approach for identification of optimal locations with in Ethiopia's wheat market to tackle soaring inflation on food price (Extended version)

Mezgebo, Taddese

Mekelle University

18 August 2009

A multivariate approach for identification of optimal locations with in Ethiopia's wheat market to tackle soaring inflation on food price
(Extended version)

By Taddese Mezgebo*
Economics Department
Collage of Business and economics
Mekelle University
Ethiopia
(2009)

Abstract

Using Johansen (1998 and 1991) vector error correction model and related extensions optimal locations where effective stabilization intervention can be done are identified. It is found that producer centers of Shashimiene and Bale Robe and Deficit market of Jimma are the locations where effective stabilization of wheat price can be done with least cost. However the national wheat market seems to have better capacity to process shocks coming from deficit markets and central markets than surplus markets. So targeting surplus markets, though effective in long run, can result on short run increase in volatility. Moreover, even though, distance could be one factor determining the inclusion of additional markets in to the rule of one price, it was not found to be the main factor in Ethiopian wheat markets. This implies that in highly imperfect market cointegration may not be solely related to distance only. Other factors like the level of market failures and the development of complementary institutions may influence the level of cointegration. Fortunately the search procedure for boundary of markets operating under rule of one price followed by earlier papers, though theoretically unsound is observed to work in Ethiopia wheat markets.

*E-mail tame_new@yahoo.com

A multivariate approach for identification of optimal locations with in Ethiopia's wheat market to tackle soaring inflation on food price
Part one
1. Introduction

After years of low and some times negative inflation Ethiopia is experiencing soaring food and general inflation. In 2000 the level of inflation in consumer price index was 6.2% and in just next two years of 2001 and 2003 it turned in to negative 5.2 and 7.2, respectively (WB 2008/9). And in general for most of the years from 1991 to 2001 the country was experiencing either very low inflation or deflation. However in recent years and mainly starting from 2006 the country is experiencing double digit inflation reaching as high as 40% in 2008 (see figure 1 below)

Figure 1 Annual inflation in Ethiopian economy
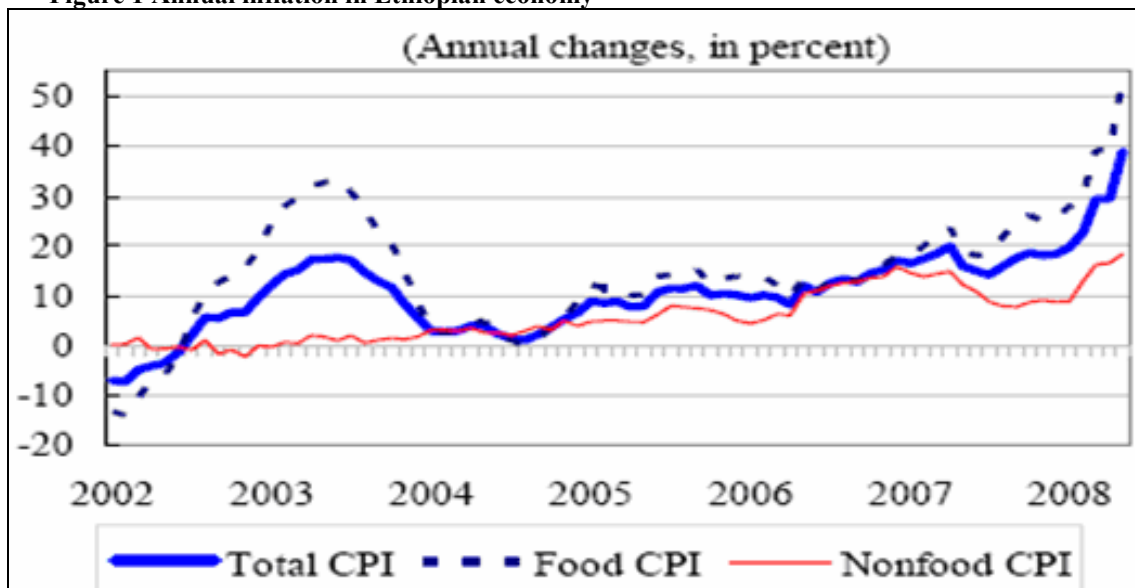
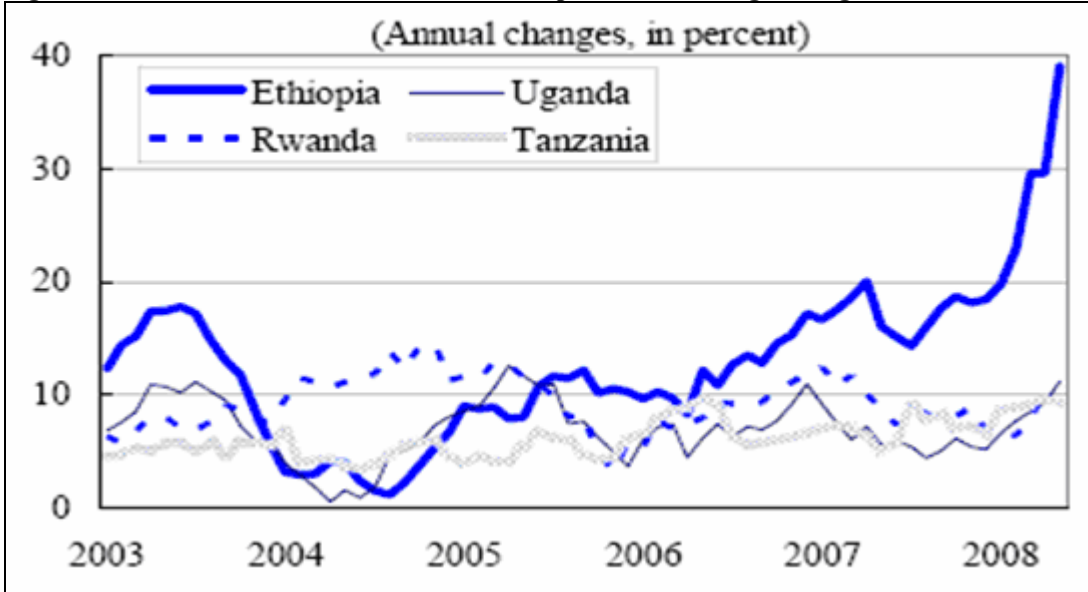**Source 1IMF(2008) IMF Country Report No. 08/259**

And as can be seen in figure 1 above this is mainly related to ever soaring food prices as the over all price index is just a shadow of the food price index. And what is puzzling is that the unprecedented level of inflation on food price is observed when all data collected by both government and international agencies is showing the country is having record level agricultural production. Moreover figure 2, below, clearly shows that the level of inflation observed in the country is much higher than the inflation observed in neighboring countries.

So the inflation observed in Ethiopia is more related to domestic economic dynamics than to global dynamics like rise in oil price, food price or any other global or regional variable. And study by Ulimwengu et al (2009) did show that there is no cointegration between domestic and international maize price supporting the view that Ethiopian food

inflation, which is the major cause of the over all inflation, is not caused by Global or regional factors but by domestic factors.

So even though understanding the source of the problem can improve the effectiveness of any intervention, consciousness does not seem to exist between stakeholders about the real domestic source of the problem and effective solution to the problem. One of the possible solutions forwarded by government is to intervene in grain markets through international purchase and distribution of white wheat in to the domestic economy.

Figure 2 Relative annual inflation between Ethiopia and three neighboring countries

(Annual changes, in percent)

Source 2 IMF(2008) IMF Country Report No. 08/259

However such intervention, if needed to be effective and efficient needs to be targeted in optimal locations where effective stabilizing intervention can be done with least possible cost. In this paper optimal locations which can be used for stabilization of wheat price are identified by using a vector error correction model (VECM) developed by Johansen (1988 and 1991) with search criteria for one common trend introduced by Gonzalo and Pitarakis (2000) with some modifications to avoid possibly inconclusive results. Moreover the short run dynamics of the market are articulated by analyzing the adjustment parameters estimated in vector error correction model and persistence profile for system level shock developed by Pesaran and Shin (1996). Additionally the markets which are having major impact on the long run common trend, which in turn is keeping the prices under rule of one price, is estimated by following Gonzalo and Granger (1995) common trend estimation methodology.

The data used for this paper is collected by European Union under price information system project and is obtained from Ethiopian Grain Trade Enterprise (EGTE). The data is extended from 1980 to 2003 However most of the data before 1996 was complied from different records and there are many missing values for many months. However for 8 whole sale markets, more or less, complete monthly data is found from 1996 to 2003. There are few missing values in some months but they are extrapolated from the data. To

extrapolate the missing value first the price is regressed on monthly dummy and year and the predicted value is used as initial value. Then given monthly nature of the data auto regressive model with 15 lags or AR(15) is fitted and the predict value is replaced for the originally missing value. And recursive estimation, prediction and replacement are done until the difference between used value and new predicted value becomes zero. The basic idea is to extrapolate the needed information from the data itself by considering the information on the lagged values of the level price.

Following this part the methodology used in this paper will be fully explained in part two. Then a unit root test based on augmented dickey fuller (ADF) test is followed by cointegration and related analysis in part three. And finally conclusion is be given in part four.

Part two
2. Econometric methodology
2.1. Introduction to econometric methodology

The main focus of the paper is to determine the long run and short run relationship that exists between wheat market prices based on vector error correction model (VECM). The level of integration of wheat markets located in different parts of the country under rule of one price is very informative in guiding stabilization efforts to optimal locations. Such optimal location selection can improve the effectiveness of stabilization policy.

If the markets in different location are highly integrated few or even one market can be used to stabilize the whole country. But if the markets are not integrated under rule of one price optimal stabilization may need simultaneous intervention in different part of the country. Moreover if there are many markets under rule of one price the market place where optimal intervention can be targeted can be identified depending on the statistical significance of the adjustment parameters, the speed of adjustment of each market to equilibrium following system wide shock and the relative importance of each market in determination of the single common trend, which is keeping the prices under rule of one price.

The prices of the same grain in different markets are expected to have an equilibrium and strong long run relationship, which can be modeled by VECM. The conventional approach to VECM is first to determine the appropriate lag by one of the few information criterions. And for given lag to use trace statistics or maximum Eigen value to determine the number of cointegration relationships found in the market (Johansen 1988 and 1991). And assuming that there are theoretical bases which can identify the cointegration relationships, it will be fruit full three step process, in which the last step is estimating and interpreting the vector error correction parameters[1].

Unfortunately for grain prices theory tale us that, if there is free flow of information and goods all market prices should be cointegrated under one common trend, unless the log of transaction costs are not stationary. If we get $n-1$ cointegration relations there is no identification problem as all markets are pair wise cointegrated and all are following a single common trend[2] (Gonzalez – Rivera and Helfand, 2001). But if the number of cointegration relationships are less than $n-1$, there will be identification problem in which neither theory nor empirical evidence will be help full. Empirical evidence mainly the cointegrating parameters identify the space spanned by the cointegrating vectors not the true cointegrating vectors. Normally theory is used to fix the restrictions needed to identify the cointegration equations. But for grain prices theory is no help in identifying cointegrated markets and nature of their relationship. So the solution is to turn the process

---

[1] Actually it is two step procedure as the rank test will be done by estimating the parameters under fully concentrated full information maximum likelihood (Hamilton, 1994). But the soft ware used in this paper stata version 9 will follow three step procedures. In which the last step will be divided in to two steps. One is rank test and the other is estimation of vector error correction model.

[2] Still out of p_n^n cointegration vectors only $n-1$ are relevant and others are redundant.

up side down as is done by Gonzalez – Rivera and Helfand (2001) on their study of Brazil rice markets and Rashid (2004) in his study of Uganda maize markets.

This non conventional approach follows the process of searching $n-1$ cointegrating prices through routine search starting from $m < n$ well connected markets to ward $n-1$ markets, which are following 1 common trend. In the Gonzalez – Rivera and Helfand (2001) paper the search was started from 10 markets which are assumed and found to be strongly cointegrated at two lags. Test for normality, ARCH effect, serial correlation and other tests However were done only for final model of 15 markets. In Rashid (2004) paper the search is started from two markets and at each stage normality test is done but test for serial correlation is not done.

In the first paper estimate of trade flow in addition to unit root test is used to determine the first 10 markets. Other markets are added sequentially given they are following 1 common trend. And it was observed that distance is an important factor on explaining, if a given market is to be part of the one long run trend or not. Moreover if the market is close to the cointegrated markets and mainly to the capital city there is high probability that it will show strong cointegration under one common trend with them. In Rashid's (2004) adaptation of the methodology the search is started from capital city and another major regional market center (Jinja). And based on their distance from the Kampala (the capital city of Uganda) other markets are added sequentially. In each sequence normality test is done and lags are added when ever necessary to achieve normality of the error terms. But test for serial correlation was not done at each stage. Unfortunately, the unit root and rank test are basically dependent on the assumption of independently distributed error terms for all sample sizes and normally, identically and independently distributed error terms for small sample sizes (Johansen 1991). So it is more logical if each search is followed by necessary testes to make sure that the error terms are white noise. In this paper test for normality, serial correlation and ARCH effect are done at each stage.

If the vector auto regressive model (VARM) of two or more markets has a shortest lag, it could possibly imply the markets under this relationship are highly integrated to one another; which in turn could imply strong codependence, compared to another group of markets. And this is in line with early models of market integration based on Ravallion (1986) bivariate auto regressive (VAR) model, which infers markets are integrated in short run if the coefficients of lagged prices are statistically equal to zero. And the shorter is the lag the more integrated the price are (Sadoulet and Janvry 1995).

But the Ravallion bivariate vector auto regressive model (VAR) has three methodological problems. First it does not consider the entire market as one structure, but it only considers two markets in isolation and this can introduce specification bias (Gonzalez – Rivera and Helfand 2001). Second it will have indignity problem since each price can granger cause the other (Sadoulet and Janvry, 1995). But the third major problem is since it uses first difference of prices it does not consider the long run relationship possibly existing between prices discovered in spatially disconnected markets. So the right way to model cointegration of prices is to use VECM which accounts for all the above short comings of the Ravallion model.

But there is a problem to this procedure when applied to cointegrated system. First in cointegration relationships we have two short term groups of parameters. These are the group of lags and the group of adjustment parameters. In bivariate VAR using first difference of prices markets with shorter lags are very closely related markets. But in VECM such conclusion is not possible since there are two groups of short run parameters. The adjustment parameters will measure the response of a market to shock initiated in given cointegrating vector. But the shocks will persist through the long memory component as reflected by the coefficients on lagged first difference. So the over all adjustment to system wide shock has to be analyzed.

Additionally the vector error correction will introduce its own identification problems. One is related to the fact that if two markets say X and Y have strong relationship at short lag with low level of persistence, it does not necessary mean any market combination with strong long run relationship need to based on them. It is possible that A and B markets can have zero or weak relationship pair wise compared to X and Y, but if C is added to A and B, the strength of A, B and C can be much stronger than X, Y and any other market. This is so because low dimension estimation of high dimension relation ships will introduce specification bias (Gonzalez – Rivera and Helfand 2001). However in both Gonzalez – Rivera and Helfand (2001) and Rashid (2004), it is assumed that any strong long run relation ship have to be based on X and Y.

But most importantly even though in Gonzalez – Rivera and Helfand (2001) the search was not found to be sensitive to the order of markets, it was not found to be true in this paper. If testes are done at each stage normality test is found to be very sensitive to slight change of order. This implies we have to search among large permutation not small combination of markets to identify markets which are operating under rule of one price.

To account for such possibility, in this paper first all possible permutation of markets are tested and the over all market dynamics is mapped by testing for one common trend using trace statistics given the error vectors are white noise. And if distance have to be the identifying parameter in order of inclusion, if the capital city need to be the center of market dynamics and if the reversal of strength can be observed or not is left to be observed from the data and it is not assumed in to the model.

The advantage of this procedure is related to the fact that it does not impose unproven assumptions in to the model. The disadvantage is that the search will be very tedious process which needs large permutation[3] of markets. In this option you need to search for large number of permutation of markets. For number of prices equal to n and maximum number lags $\left(L_i \right)$ considered in i permutation of markets[4] the total number permutation of markets to be tested for each model is equal to

$$\left(p_2^n \times L_2 \right) + \left(p_3^n \times L_3 \right) + \left(p_4^n \times L_4 \right) + \ldots\ldots + \left(p_n^n \times L_n \right)$$

---

[3] Permutations than combinations have to be used since normality is order sensitive.
[4] For the 8 wheat prices traded in different markets it took around 120 hours of computer time on computer with 1.79 GHZ processor.

In this study given small sample size of 96 observations of monthly price dated from 1996 to 2003, it is pushed for strict normality and lack of serial correlation. So unless the null of normality and independence can not be rejected up to 10% level of significance, the hypothesis of correct market order is rejected. If the market permutation is having independently and normally distributed error vector, ARCH test is also done to measure the significance of time varying volatility.

For VEC the first step is to find prices which are generated by data generating process with unit root. So let's start from the theoretical exposition of the ADF unit root test and related F version tests before vector error correction model is developed. In this paper unless specifically stated to contrary, there is high dependence on theoretical exposition, of different models developed by many authors, given by Hamilton (1994).

2.2. <u>Theoretical base for Augmented Dickey -Fuller Test (ADF) and related F kind of tests for unit root</u>

Let's take a log price of one commodity in given market (P) in given time (t) by a scalar P_t. The price can be presented by following Auto Regressive Integrated Moving Average (ARIMA) representation with p auto regressive terms, I integrated term[5] and q moving average terms. Formally $ARIMA(p,I,q)$ of price with trend is given by

$$P_t = \alpha * + \beta * t + \theta_1 P_{t-1} + \theta_2 P_{t-2} + ... \ ... + \theta_p P_{t-p} + \varepsilon_t + \lambda_1 \varepsilon_{t-1} + \lambda_2 \varepsilon_{t-2} + ... \ ... + \lambda_q \varepsilon_{t-q} ....1$$

This can be presented concisely using lag operator as

$$P_t = \alpha * + \beta * t + \theta_1 L P_t + \theta_2 L^2 P_t + ... \ ... + \theta_p L^p P_t + \varepsilon_t + \lambda_1 L \varepsilon_t + \lambda_2 L^2 \varepsilon_t + ... \ ... + \lambda_q L^q \varepsilon_t$$

$$\left(1 - \theta_1 L - \theta_2 L^2 - ... \ ... - \theta_p L^p\right) P_t = \alpha * + \beta * t + \left(1 + \lambda_1 L + \lambda_2 L^2 + ... \ ... + \lambda_q L^q\right) \varepsilon_t$$

$$\theta(L) P_t = \alpha * + \beta * t + \lambda(L) \varepsilon_t \ ....2$$

All roots of $\lambda(L)$ are assumed to have module greater than 1 or Eigen value less than 1, means $\lambda(L)$ is invertible. So equation 2 can be represented by following infinitive order auto regressive term or $AR(\infty)$,

$$\left[\lambda(L)\right]^{-1} \theta(L) P_t = \left[\lambda(L)\right]^{-1} \alpha * + \left[\lambda(L)\right]^{-1} \beta * t + \varepsilon_t$$

$$\phi(L) P_t = \alpha + \beta t + \varepsilon_t$$

$$\left(1 - \phi_1 L - \phi_2 L^2 ... \ ... - \phi_\infty L^\infty\right) P_t = \alpha + \beta t + \varepsilon_t \ ....3$$

For $\left[\lambda(L)\right]^{-1} \theta(L) = \phi(L)$, $\alpha = \left[\lambda(L)\right]^{-1} \alpha *$ and $\beta = \left[\lambda(L)\right]^{-1} \beta *$. But given lack of an infinitive data and to allow for higher degree freedom, the infinitive auto regressive lag is approximated by $p + s = m$ lag. The size of m is chosen by selecting a practical lag length at which the highest information is incorporated in to the model at lowest cost in

---

[5] or I unit roots

terms of degree of freedom. Means equation 3 is approximated by auto regressive with m lag or $AR(m)$. Formally

$$\left(1 - \phi_1 L - \phi_2 L^2 \dots \dots - \phi_m L^m\right) P_t = \alpha + \beta t + \varepsilon_t$$

$$\phi(L) P_t = \alpha + \beta t + \varepsilon_t \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots.4$$

For $\phi(L)$ is redefined as $\left(1 - \phi_1 L - \phi_2 L^2 \dots \dots - \phi_m L^m\right)$ in equation 4. So any general $ARIMA(p,I,q)$ model on price can be presented by $AR(m)$ model with appropriate selection of the necessary lag order or m. If there is just one unit root in equation 4 one Eigen value of the lag operator $\phi(L)$ will be one but other Eigen values of $\phi(L)$ will have module less than one.

Formally for $\left(1 - \phi_1 z - \phi_2 z^2 \dots \dots - \phi_m z^m\right) = \left(1 - \phi_1 z\right)\left(1 - \phi_2 z^2 - \phi_3 z^3 \dots\dots - \phi_{m-1} z^{m-1}\right) = 0$, $\phi_1 = 1$ and $\left|\phi_j\right| < 1$ for all $j \neq 1$. So at $z = 1$, if there is just one unit root it will follow that

$$1 - \phi_1 - \phi_2 - \phi_3 - \dots - \phi_m = 0 \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots.5$$

Keeping this in mind lets go back to equation 4, above. For $\rho = \left(\phi_1 + \phi_2 + \phi_3 + \dots + \phi_m\right)$ and $\varsigma_i = -\left(\phi_{i+1} + \phi_{i+2} + \dots + \phi_m\right)$ the following is true about the lag operator at equation 4

$$\left(1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_m L^m\right) = \left(1 - \rho L\right) - \left(\varsigma_1 L + \varsigma_2 L^2 + \dots + \varsigma_{m-1} L^{m-1}\right)\left(1 - L\right) \dots\dots\dots\dots.6$$

So by replacing equation 6 in to equation 4 we have the following representation of the auto regressive term

$$\left[\left(1 - \rho L\right) - \left(\varsigma_1 L + \varsigma_2 L^2 + \dots + \varsigma_{m-1} L^{m-1}\right)\left(1 - L\right)\right] P_t = \alpha + \beta t + \varepsilon_t$$

$$\left(P_t - \rho L P_t\right) - \left(\varsigma_1 L + \varsigma_2 L^2 + \dots + \varsigma_{m-1} L^{m-1}\right)\Delta P_t = \alpha + \beta t + \varepsilon_t$$

$$P_t = \alpha + \beta t + \varsigma_1 L \Delta P_t + \varsigma_2 L^2 \Delta P_t + \dots + \varsigma_{m-1} L^{m-1}\Delta P_t + \rho L P + \varepsilon_t$$

$$P_t = \alpha_t + \beta t + \varsigma_1 \Delta P_{t-1} + \varsigma_2 \Delta P_{t-2} + \dots + \varsigma_{m-1} \Delta P_{t-m+1} + \rho P_{t-1} + \varepsilon_t$$

$$P_t - P_{t-1} = \alpha_t + \beta t + \varsigma_1 \Delta P_{t-1} + \varsigma_2 \Delta P_{t-2} + \dots + \varsigma_{m-1} \Delta P_{t-m+1} + \rho P_{t-1} - P_{t-1} + \varepsilon_t$$

$$\Delta P_t = \alpha + \beta t + \varsigma_1 \Delta P_{t-1} + \varsigma_2 \Delta P_{t-2} + \dots + \varsigma_{m-1} \Delta P_{t-m+1} + \left(\rho - 1\right) P_{t-1} + \varepsilon_t$$

$$\Delta P_t = \alpha + \beta t + \varsigma_1 \Delta P_{t-1} + \varsigma_2 \Delta P_{t-2} + \dots + \varsigma_{m-1} \Delta P_{t-m+1} - \left(1 - \rho\right) P_{t-1} + \varepsilon_t$$

$$\Delta P_t = \alpha_t + \beta t + \varsigma_1 \Delta P_{t-1} + \varsigma_2 \Delta P_{t-2} + \dots + \varsigma_{m-1} \Delta P_{t-m+1} - \lambda P_{t-1} + \varepsilon_t \dots\dots\dots\dots\dots\dots.7$$

But from equation 5 the condition for having unit root is satisfied if $1 - \rho = 1 - \phi_1 - \phi_2 - \dots - \phi_m = 0$ or if $\rho = 1$. This means the test for unit root is a test to see if λ is statically equal to zero or not. This is the logic behind the famous ADF test developed by Dickey and Fuller (1979). The test can be done by using either the ADF test with out drift (t_{df}), with drift (t_α) or with deterministic time trend (t_T). What will change is the distribution to be compared against not the calculation of the test statistics.

For σ_λ standard error of λ the ADF statistics are calculated as

$$t_{df} = t_\alpha = t_T = \frac{\lambda}{\sigma_\lambda} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 8$$

Seen from asymptotic theory, if the real data generating process is with out drift and equation number 7 is fitted by dropping the drift term $(\alpha_t = 0)$ and the trend term $(\beta = 0)$ then t_{df} version of ADF have to be used. But if the model with drift is used when the true data generating process is with out drift, t_α version ADF have to be used. If the real data generating process is with drift and equation 7 is used with the drift term, then the distribution is asymptotically Gaussian so normal distribution can be used. And if equation 7 is estimated with out any restriction, then t_T version of ADF test has to be used (Dickey and Fuller 1979).

But there are two basic problems observed in practice. First the true data generating process is unknown so one can use one distribution among others to justify what ever conclusion that he/she wants to make (Peterson, 2000). But most importantly for sample size of 96 months depending on asymptotic theory to justify any conclusion is not sound. And Monte Carlo experiments did show that in small sample sizes not only the existence of drift term but also the value of the drift term used in Monte Carlos study is observed to make a huge difference on the critical values (Ibid). But unless the ADF test is made we can't know if the use of normal distribution is justifiable. And unless we know it is justified to use normal distribution we can't test for the true value of the drift. And unless we know the true value of the drift we can't make use of ADF test. This is a circular chicken and egg problem most manifested in small sample. If the trend term is included the value of drift is not important as it is orthogonal with λ (Dickey and Fuller 1979). But still the critical values are dependent on the coefficient of the trend term (β) used or assumed in data generating process mainly for small samples (Peterson 2000). Additionally using t_T when the true data generating process is with out trend will lead to weak power of the test (Hamilton 1994 and Peterson 2000).

In econometrics "which is science and art" (Green 2003) there is way around the problem by using F- kind of tests termed in this paper as Φ_1 and Φ_3. The first one will check for joint significance of the drift (α) and the unit root term (λ) under the null that both are zero. If it is accepted that unit root is accepted to be found in the data with zero drift. If rejected However there are three possibilities (Peterson 2000).

1. The first option is that the data is stationary around zero or the drift term (α) is zero but the unit root term (λ) is different from zero. Means the mean of the data is zero. But in such case the use of Φ_1 is not justifiable.
2. The second option is that (α) is different from zero and (λ) is zero means the data follows unit root around drift. Means there is stochastic trend in the data but in such

case we have to use Φ_3 than Φ_1.

3. Third possibility is that both the drift term (α) and the unit root term (λ) are different from zero. Means it is stationery around drift term. And this is the most logical conclusion. Notice that if the data is stationary α will be related to none zero mean but if the data is following unit root α will represent stochastic trend in the data.

The basic logic is that Φ_1 have to be used only if there is none zero mean on the data as observed from visual inspection but there must be neither random nor deterministic trend on it. And it is use full if we are not sure if the none zero mean is caused by long memory of error terms related to random walk around zero (Null hypothesis) or by drift in stationary series (Alternative hypothesis). Means if we are using Φ_1 the data must have none zero average value and must not have any trend. This will rule out alternative 1 and 2. This is why econometrics is both science and art. You use mathematical model (science) as far as it can take you. When practical problems can't be solved mathematically, However, you use creative thinking (art) to go around the problem.

The use of Φ_3 will be justified only if there is visible trend in the level data. And Φ_3 checks for joint significance of the unit root coefficient (λ) and the trend coefficient (β) under the null hypothesis that both are equal to zero. If it is accepted then the data have unit root around drift or is having stochastic trend. But if it is rejected there are three possibilities (Ibid)

1. The trend coefficient is zero $(\beta = 0)$ and there is no unit root $(\lambda \neq 0)$. Means the data is stationary around drift term or it have none zero average but no trend. But the use of Φ_3 is justified only if there is visible trend in the data.

2. The data have unit root $(\lambda = 0)$ around deterministic trend $(\beta \neq 0)$. But this implies there is random and deterministic trend in the data. This can happen only if the level data have quadratic trend. And most data in economics are linearly trended, at best.

3. Both the trend coefficient (β) and the unit root coefficient (λ) are different from zero. Means the data is stationary around deterministic trend. And this is the most logical conclusion unless there is quadratic trend on the level data. Notice again that the trend coefficient will generate quadratic trend in level data, if there is unit root. But it will generate linear trend, if the data is stationery.

If the data have visible trend and the use of Φ_3 justified the rejection of the null implies the data is stationary around deterministic trend. And acceptance of the null implies that the data have unit root around drift. What it means is that we observe a linear trend in the data but we are not sure, if it is caused by random walk with drift (null hypothesis) or trend related to deterministic time with stationary series (alternative hypothesis). So we use Φ_3 to conclude one way or another. Means to determine, if the observed trend is stochastic in nature related to unit root around drift or deterministic in nature related to

stationary series around trend. If there is quadratic trend in the level data it is possible unit root can be found with linear trend (option 2). But this is rare in economic data but for curiosity visual inspection is needed to make sure that there is no quadratic trend in the level data.

Once these results are conformed then the asymptotic theory based on ADF T version test can be used to farther consolidate the conclusion. So the test for unit root will be done from general to specific as advised by Peterson (2000). But given the lack of rigid formula for unit root and the weak power of Φ_1 and Φ_3 it is important to accept any conclusion as pragmatic approximation to the true distribution than the true distribution per se. See Hamilton (1994) and Peterson (2000) for farther discussion in this issue.

In this paper both visual information and test results will be given to back the conclusions of the study. But once the existence of unit root is conformed the next step is to model the relationship between prices. Given prices are expected to have strong long run relationship, as conformed by economic theory and wide empirical evidence, the model chosen in this paper is Vector Error Correction (VEC) representation of Vector Auto Regressive Integrated Moving Average model (VARIMA) as given by Granger representation theorem (Engel and Granger 1987). The estimation will be done by using Johansen (1998 and 1991) full information maximum likelihood estimation technique. The granger representation theorem and the Johansen full information maximum likelihood estimation technique will be developed below. As was stated above most of the model specifications used in this paper, unless specifically stated to contrary, are adapted to vector of prices from general theoretical exposition of time series models given by Hamilton (1994)

2.3. Theoretical base of Cointegration analysis based on vector error correction model (VECM)

The log price (P) of a commodity in given market (i) can be presented by variable P_i. And all n log prices in n locations can be presented by n dimensional vector $\mathbf{P}$.

$$\mathbf{P}_t' = \begin{bmatrix} P_1 & P_2 & . & . & . & P_n \end{bmatrix} \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots 9$$

And let's assume the vector $\mathbf{P}_t$ or the vector of log prices at time t, can be presented by the following Vector Autoregressive Integrated Moving Average form or VARIMA$(p,\, I,\, q)$ with trend

$$\mathbf{P}_t = \boldsymbol{\delta} + \boldsymbol{\pi}\mathbf{t} + \boldsymbol{\lambda}_1\mathbf{P}_{t-1} + \boldsymbol{\lambda}_2\mathbf{P}_{t-2}... + \boldsymbol{\lambda}_p\mathbf{P}_{t-p} + \boldsymbol{\varepsilon}_t + \boldsymbol{\gamma}_1\boldsymbol{\varepsilon}_{t-1} + \boldsymbol{\gamma}_2\boldsymbol{\varepsilon}_{t-2}... + \boldsymbol{\gamma}_q\boldsymbol{\varepsilon}_{t-q} \dots\dots\dots\dots 10$$

Vector $\boldsymbol{\varepsilon}_t$ is white noise with variance covariance matrix $\boldsymbol{\Omega}$. In which $E(\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_\tau') = 0$ for $t \neq \tau$ and $E(\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_\tau') = \boldsymbol{\Omega}$ for $t = \tau$. Where $\boldsymbol{\lambda}_i$ and $\boldsymbol{\gamma}_i$ are $(n \times n)$ matrixes representing the vector autoregressive and moving average coefficients, respectively. The trend term $(\mathbf{t})$ is related to time and its coefficient $(\boldsymbol{\pi})$ is n dimension vector. Since the number of unit

roots is not know yet, lets ignore the I or integrated part of the above equation 10. If rearranged by using lag operator equation 10 will be

$$P_t - \lambda_1 P_{t-1} - \lambda_2 P_{t-2}... - \lambda_p P_{t-p} = \delta + \pi t + \varepsilon_t + \gamma_1 \varepsilon_{t-1} + \gamma_2 \varepsilon_{t-2}... + \gamma_q \varepsilon_{t-q}$$

$$P_t - \lambda_1 L P_t - \lambda_2 L^2 P_t... - \lambda_p L^p P_t = \delta + \pi t + \varepsilon_t + \gamma_1 L \varepsilon_t + \gamma_2 L^2 \varepsilon_t... + \gamma_q L^q \varepsilon_t$$

$$\Theta(L) P_t = \delta + \pi t + \Upsilon(L) \varepsilon_t \quad\text{...11}$$

Lets assuming there is only one unit root for each log price in $\Theta(L)$ or in other words the roots of $\Theta(Z)$ have one module equal to 1, while the module of the rest of the roots are out of the unit circle. If we define $\Theta(L) = (I-L)\Theta^*(L)$ then $\Theta^*(L)$ is invertible with all its roots having module greater than one. So

$$(I-L)\Theta^*(L) P_t = \delta + \pi t + \Upsilon(L)\varepsilon_t$$

$$(I-L) P_t = \left[\Theta^*(L)\right]^{-1}\delta + \left[\Theta^*(L)\right]^{-1}\pi t + \left[\Theta^*(L)\right]^{-1}\Upsilon(L)\varepsilon_t$$

$$(I-L) P_t = \alpha + \beta t + \Psi(L)\varepsilon_t \quad\text{...12}$$

This is known as wold representation of P_t, in which $\left[\Theta^*(L)\right]^{-1}\Upsilon(L) = \sum_{i=0}^{\infty}\Psi_i\varepsilon_{t-i} = \Psi(L)$, $\left[\Theta^*(L)\right]^{-1}\delta = \alpha$ and $\left[\Theta^*(L)\right]^{-1}\pi = \beta$. Moreover for $\Psi(L)\varepsilon_t = \mu_t$ equation 12 will simplify to

$$\Delta P_t = \alpha + \beta t + \mu_t \quad\text{...13}$$

Means the change in price is equal to drift vector(α), trend effect (βt) and random shock or innovation vector (μ_t) which take a moving average form with infinitive lag. So for

$$P_t = \Delta P_t + \Delta P_{t-1} + \Delta P_{t-2} \text{.......................} + \Delta P_1 + P_0$$

$$P_t = t\alpha + \beta \sum_{i=1}^{t} i + \mu_t + \mu_{t-1} + \mu_{t-2} \text{.....................} + \mu_1 + P_0$$

$$P_t = t\alpha + \beta \frac{t(t+1)}{2} + \mu_t + \mu_{t-1} + \mu_{t-2} \text{.....................} + \mu_1 + P_0$$

$$P_t = t\alpha + \beta \frac{(t^2 + t)}{2} + \sum_{i=1}^{t} \mu_i + P_0 \quad\text{...14}$$

[6]Let's concentrate on the sum of the moving average terms given as

$$\mu_t = \sum_{i=0}^{\infty} \Psi_i \varepsilon_{t-i} = \varepsilon_t + \Psi_1 \varepsilon_{t-1} + \Psi_2 \varepsilon_{t-2} + \Psi_3 \varepsilon_{t-3} \quad \dots\dots$$

$$\sum_{i=1}^{t} \mu_i = \begin{bmatrix} \left(\varepsilon_t + \Psi_1 \varepsilon_{t-1} + \Psi_2 \varepsilon_{t-2} \dots\right) + \left(\varepsilon_{t-1} + \Psi_1 \varepsilon_{t-2} + \Psi_2 \varepsilon_{t-3} \dots\right) \\ + \left(\varepsilon_{t-2} + \Psi_1 \varepsilon_{t-3} + \Psi_2 \varepsilon_{t-4} \dots\right) \dots + \left(\varepsilon_1 + \Psi_1 \varepsilon_{-1} + \Psi_2 \varepsilon_{-2} \dots\right) \end{bmatrix}$$

$$\sum_{i=1}^{t} \mu_i = \begin{bmatrix} \varepsilon_t + \left(\varepsilon_{t-1} + \Psi_1 \varepsilon_{t-1}\right) + \left(\varepsilon_{t-2} + \Psi_1 \varepsilon_{t-2} + \Psi_2 \varepsilon_{t-2}\right) + \\ \left(\varepsilon_{t-3} + \Psi_1 \varepsilon_{t-3} + \Psi_2 \varepsilon_{t-3} + \Psi_3 \varepsilon_{t-3}\right) + \dots \\ + \left(\varepsilon_1 + \Psi_1 \varepsilon_1 + \Psi_2 \varepsilon_1 + \Psi_3 \varepsilon_1 \dots \Psi_{t-1} \varepsilon_1\right) + \left(\varepsilon_0 + \Psi_1 \varepsilon_0 + \Psi_2 \varepsilon_0 + \Psi_3 \varepsilon_0 \dots \Psi_t \varepsilon_0\right) \\ + \left(\varepsilon_{-1} + \Psi_1 \varepsilon_{-1} + \Psi_2 \varepsilon_{-1} + \Psi_3 \varepsilon_{-1} \dots \Psi_{t+1} \varepsilon_{-1}\right) + \dots\dots\dots\dots\dots \end{bmatrix}$$

$$\sum_{i=1}^{t} \mu_i = \begin{bmatrix} \left(I + \Psi_1 + \Psi_2 + \Psi_3 \dots\right)\varepsilon_t + \left(\Psi_1 + \Psi_2 + \Psi_3 \dots\right)\varepsilon_t + \\ \left(I + \Psi_1 + \Psi_2 + \Psi_3 \dots\right)\varepsilon_{t-1} + \left(\Psi_2 + \Psi_3 + \Psi_4 \dots\right)\varepsilon_{t-1} + \\ \left(I + \Psi_1 + \Psi_2 + \Psi_3 \dots\right)\varepsilon_{t-2} + \left(\Psi_3 + \Psi_4 + \Psi_5 \dots\right)\varepsilon_{t-2} + \dots\dots\dots\dots + \\ \left(I + \Psi_1 + \Psi_2 + \Psi_3 \dots\right)\varepsilon_1 + \left(\Psi_t + \Psi_{t+1} + \Psi_{t+2} \dots\right)\varepsilon_1 + \\ \left(\Psi_1 + \Psi_2 + \Psi_3 + \Psi_4 \dots\right)\varepsilon_0 + \left(\Psi_{t+1} + \Psi_{t+2} + \Psi_{t+3} \dots\right)\varepsilon_0 + \\ \left(\Psi_2 + \Psi_3 + \Psi_4 + \Psi_5 \dots\right)\varepsilon_{-1} + \left(\Psi_{t+2} + \Psi_{t+3} + \Psi_{t+4} \dots\right)\varepsilon_{-1} + \dots\dots \end{bmatrix} \dots\dots\dots\dots 15$$

Taking the upper left for ε related to period of 1 to t will give us the following random walk vector $\Psi(1)\sum_{t=1}^{t}\varepsilon_t$, for

$$\Psi(1)\sum_{t=1}^{t}\varepsilon_t = \left(I + \Psi_1 + \Psi_2 \dots\right)\varepsilon_T + \left(I + \Psi_1 + \Psi_2 \dots\right)\varepsilon_{T-1} + \dots + \left(I + \Psi_1 + \Psi_2 \dots\right)\varepsilon_1.$$

Taking all the values in the right up for all time periods from t up to negative infinitive will give us a stationery vector η_t for

$$\eta_t = -\left(\Psi_1 + \Psi_2 + \Psi_3 \dots\right)\varepsilon_t - \left(\Psi_2 + \Psi_3 + \Psi_4 \dots\right)\varepsilon_{t-1} - \left(\Psi_3 + \Psi_4 + \Psi_5 \dots\right)\varepsilon_{t-2} + \dots\dots\dots\dots$$

And taking the lower left vectors related to time period of 0 to negative infinitive will give us initial information contained in η_0 or

$$\eta_0 = -\left(\Psi_1 + \Psi_2 + \Psi_3 + \Psi_4 \dots\right)\varepsilon_0 - \left(\Psi_2 + \Psi_3 + \Psi_4 + \Psi_5 \dots\right)\varepsilon_{-1} \dots\dots\dots$$

So the sum of the moving average terms given in equation 14 is equal to

---

[6] This explains why in unit root tests drift with unit root is associated with trend on level and trend with unit root is associated with quadratic trend in level data. You can apply the logic to single price and you can get the same representation as above.

$$\sum_{i=1}^{t} \mu_i = \Psi(1) \sum_{i=1}^{t} \varepsilon_i + \eta_t - \eta_0 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots16$$

By replacing equation 16 in to equation 14, we can get the following moving average representation

$$P_t = t\alpha + \frac{1}{2}\beta(t^2 + t) + \Psi(1)\sum_{i=1}^{t}\varepsilon_i + \eta_t + (P_0 - \eta_0) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots17$$

So P_t is represented in terms of the sum of trend times drift term $(t\alpha)$, a random walk variable $\left(\Psi(1)\sum_{i=1}^{t}\varepsilon_i\right)$, quadratic deterministic trend $(t^2 + t)$ and two stationary variables. One related to initial condition $(P_0 - \eta_0)$ and another related to a stationary process (η_t).

So if there is a $(n \times h)$ matrix $\mathbf{A}$ for which $\mathbf{A}P_t$ is stationary or $\mathbf{A}$ is cointegrating matrix of P_t then both the left and the right side of the following equation have to be stationary

$$\mathbf{A}P_t = t\mathbf{A}\alpha + \frac{1}{2}\mathbf{A}\beta(t^2 + t) + \mathbf{A}\Psi(1)\sum_{i=1}^{t}\varepsilon_i + \mathbf{A}\eta_t + \mathbf{A}(P_0 - \eta_0) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots18$$

Given the space spanned by stationary series will produce stationary series the two last vectors are stationary by definition. So this implies for cointegrating matrix $\mathbf{A}$ the following must be true

$$\mathbf{A}\alpha = 0 \,, \; \mathbf{A}\beta = 0 \; \text{and} \; \mathbf{A}\Psi(1) = 0 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots19$$

If the above three conditions are not attained then $\mathbf{A}$ is not a cointegrating matrix. Now let's bring back equation 11 below again

$$\Theta(\mathbf{L})P_t = \delta + \pi t + \Upsilon(\mathbf{L})\varepsilon_t \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots11$$

Since $\Upsilon(\mathbf{L})$ is invertible or it have Eigen values with module less than one. Equation 11 can be expressed in vector auto regressive form with infinitive autoregressive terms

$$\left[\Upsilon(\mathbf{L})\right]^{-1}\Theta(\mathbf{L})P_t = \left[\Upsilon(\mathbf{L})\right]^{-1}\delta + \left[\Upsilon(\mathbf{L})\right]^{-1}\pi t + \varepsilon_t$$

$$\Phi(L)P_t = \delta^* + \pi^* t + \varepsilon_t \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots20$$

For $\left[\Upsilon(\mathbf{L})\right]^{-1}\delta = \delta^*$, $\left[\Upsilon(\mathbf{L})\right]^{-1}\pi = \pi^*$ and $\left[\Upsilon(\mathbf{L})\right]^{-1}\Theta(\mathbf{L}) = \Phi(\mathbf{L})$. $\Phi(\mathbf{L})$ represents an infinitive lag operator or $\Phi(\mathbf{L}) = \mathbf{I} - \Phi_1 L^1 - \Phi_2 L^2 - \Phi_3 L^3 - \Phi_4 L^4 \ldots - \Phi_\infty L^\infty$. So the Vector Auto Regressive Moving Average presentation of price can be easily represented by Vector Auto Regressive with infinite lag or $VAR(\infty)$.

But again $\mathbf{P}_T$ also have Wold presentation as its first difference can be expressed as infinitive lag moving average term. This is shown in equation 12 and is replicated below

$$\left(\mathbf{I} - \mathbf{L}\right)\mathbf{P}_t = \boldsymbol{\alpha} + \boldsymbol{\beta}t + \boldsymbol{\Psi}\left(\mathbf{L}\right)\boldsymbol{\varepsilon}_t \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\text{......}12$$

And if both sides of equation 12 are multiplied by $\boldsymbol{\Phi}(L)$ it will give us

$$\left(\mathbf{I} - \mathbf{L}\right)\boldsymbol{\Phi}\left(L\right)\mathbf{P}_t = \boldsymbol{\Phi}\left(L\right)\boldsymbol{\alpha} + \boldsymbol{\Phi}\left(L\right)\boldsymbol{\beta}t + \boldsymbol{\Phi}\left(L\right)\boldsymbol{\Psi}\left(\mathbf{L}\right)\boldsymbol{\varepsilon}_t \dots\dots\dots\dots\dots\dots\dots 21$$

And replacing equation 20 in to 21 will result on

$$\left(\mathbf{I} - \mathbf{L}\right)\left(\boldsymbol{\delta}^* + \boldsymbol{\pi}^* t + \boldsymbol{\varepsilon}_t\right) = \boldsymbol{\Phi}\left(L\right)\boldsymbol{\alpha} + \boldsymbol{\Phi}\left(L\right)\boldsymbol{\beta}t + \boldsymbol{\Phi}\left(L\right)\boldsymbol{\Psi}\left(\mathbf{L}\right)\boldsymbol{\varepsilon}_t$$
$$\boldsymbol{\pi}^* + \left(\mathbf{I} - \mathbf{L}\right)\boldsymbol{\varepsilon}_t = \boldsymbol{\Phi}\left(L\right)\boldsymbol{\alpha} + \boldsymbol{\Phi}\left(L\right)\boldsymbol{\beta}t + \boldsymbol{\Phi}\left(L\right)\boldsymbol{\Psi}\left(\mathbf{L}\right)\boldsymbol{\varepsilon}_t \dots\dots\dots\dots\dots\dots\dots 22$$

So for all values of L the left and right side of equation 22 are equal. Mainly for $L = \mathbf{I}$ equation 22 implies that

$$\boldsymbol{\pi}^* + \left(\mathbf{I} - \mathbf{I}\right)\boldsymbol{\varepsilon}_t = \boldsymbol{\Phi}\left(1\right)\boldsymbol{\alpha} + \boldsymbol{\Phi}\left(1\right)\boldsymbol{\beta}t + \boldsymbol{\Phi}\left(1\right)\boldsymbol{\Psi}\left(1\right)\boldsymbol{\varepsilon}_t$$
$$\boldsymbol{\pi}^* = \boldsymbol{\Phi}\left(1\right)\boldsymbol{\alpha} + \boldsymbol{\Phi}\left(1\right)\boldsymbol{\beta}t + \boldsymbol{\Phi}\left(1\right)\boldsymbol{\Psi}\left(1\right)\boldsymbol{\varepsilon}_t \dots\dots\dots\dots\dots\dots\dots\dots\dots 23$$

And these can only happen if

$$\boldsymbol{\Phi}(1)\boldsymbol{\Psi}(1) = \mathbf{0} \text{ and } \boldsymbol{\pi}^* = \boldsymbol{\Phi}(1)\boldsymbol{\alpha} + \boldsymbol{\Phi}(1)\boldsymbol{\beta} \dots\dots\dots\dots\dots\dots\dots\dots\dots 24$$

Taking the first one or $\boldsymbol{\Phi}(1)\boldsymbol{\Psi}(1) = \mathbf{0}$ it means $\boldsymbol{\Phi}(1)$ is cointegrating vector of $\boldsymbol{\Psi}(1)$. But from equation 18 and 19 the basis of the cointegrating vector of $\boldsymbol{\Psi}(1)$ is given by $\mathbf{A}$. Means $\boldsymbol{\Phi}(1)$ and $\mathbf{A}$ did span the same vector space. So for given $(n \times h)$ matrix $\mathbf{B}$ one can be expressed as linear combination of the other. Formally

$$\boldsymbol{\Phi}(1) = \mathbf{BA'} \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots 25$$

In vector error correction $\mathbf{A}$ is the cointegration equations which define the long run relationships between prices. Prices are not expected to be in line with their long run common trend all the time. So there will be random deviation in short run due to random factors but any deviation from this equilibrium will be adjusted in time. And this adjustment process is defined by the adjustment parameters $\mathbf{B}$ [7]. Now let's replicate equation 20 below.

$$\boldsymbol{\Phi}\left(L\right)\mathbf{P}_t = \boldsymbol{\delta}^* + \boldsymbol{\pi}^* t + \boldsymbol{\varepsilon}_t \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\text{...}20$$

---

[7] In vector error correction only the space spanned by $\mathbf{A}$ and $\mathbf{B}$ can be identified but not the specific matrix $\mathbf{A}$ and $\mathbf{B}$. Either theory or some search mechanism have to be used to identify $\mathbf{A}$ and $\mathbf{B}$ from their liner combination estimated by Johnson Method to be discussed below.

For $\Phi(\mathbf{L}) = \mathbf{I} - \Phi_1 L^1 - \Phi_2 L^2 - \Phi_3 L^3 - \Phi_4 L^4 ..... - \Phi_\infty L^\infty$. But in practice Vector auto regressive up to some acceptable large lag equal to $P = p + s$ is used to approximate equation 20. Behind acceptable lags the additional information used by adding lag will be over weighted by the loss of degree of freedom. So it is very parsimonious in terms of estimation efficiency and pragmatism to use the following approximation in to equation 20

$$\left(\mathbf{I} - \Phi_1 L - \Phi_2 L^2 - \Phi_3 L^3 - ... \ ... - \Phi_P L^P \right) \mathbf{P}_t = \delta^* + \pi^* t + \varepsilon_t \quad\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots..26$$

But for $\rho = \left(\Phi_1 + \Phi_2 + \Phi_3 + ... + \Phi_P \right)$ and $\zeta_i = -\left(\Phi_{i+1} + \Phi_{i+2} + ... + \Phi_P \right)$ the following fact did hold

$$\left(\mathbf{I} - \Phi_1 L - \Phi_2 L^2 - ... - \Phi_P L^P \right) = \left(\mathbf{I} - \rho L \right) - \left(\zeta_1 L + \zeta_2 L^2 + ... + \zeta_{P-1} L^{P-1} \right)\left(\mathbf{I} - L \right) \quad\ldots\ldots\ldots\ldots..27$$

And if equation 27 is replaced in to equation 26, it will give us the following representation

$$\left[\left(\mathbf{I} - \rho L \right) - \left(\zeta_1 L + \zeta_2 L^2 + ... + \zeta_{P-1} L^{P-1} \right)\left(\mathbf{I} - L \right) \right] \mathbf{P}_t = \delta^* + \pi^* t + \varepsilon_t$$

$$\left(\mathbf{I} - \rho L \right) \mathbf{P}_t - \left(\zeta_1 L + \zeta_2 L^2 + ... + \zeta_{P-1} L^{P-1} \right) \Delta \mathbf{P}_t = \delta^* + \pi^* t + \varepsilon_t$$

$$\left(\mathbf{P}_t - \rho L \mathbf{P}_t \right) = \delta^* + \pi^* t + \zeta_1 L \Delta \mathbf{P}_t + \zeta_2 L^2 \Delta \mathbf{P}_t + ... + \zeta_{P-1} L^{P-1} \Delta \mathbf{P}_t + \varepsilon_t$$

$$\mathbf{P}_T - \rho \mathbf{P}_{t-1} = \delta^* + \pi^* t + \zeta_1 \Delta \mathbf{P}_{t-1} + \zeta_2 \Delta \mathbf{P}_{t-2} + ... + \zeta_{P-1} \Delta \mathbf{P}_{t-p+1} + \varepsilon_t$$

$$\mathbf{P}_t = \delta^* + \pi^* t + \zeta_1 \Delta \mathbf{P}_{t-1} + \zeta_2 \Delta \mathbf{P}_{t-2} + ... + \zeta_{P-1} \Delta \mathbf{P}_{t-p+1} + \rho \mathbf{P}_{t-1} + \varepsilon_t$$

$$\left(\mathbf{P}_t - \mathbf{P}_{t-1} \right) = \delta^* + \pi^* t + \zeta_1 \Delta \mathbf{P}_{t-1} + \zeta_2 \Delta \mathbf{P}_{t-2} + ... + \zeta_{P-1} \Delta \mathbf{P}_{t-p+1} + \left(\rho \mathbf{P}_{t-1} - \mathbf{P}_{t-1} \right) + \varepsilon_t$$

$$\Delta \mathbf{P}_t = \delta^* + \pi^* t + \zeta_1 \Delta \mathbf{P}_{t-1} + \zeta_2 \Delta \mathbf{P}_{t-2} + ... + \zeta_{P-1} \Delta \mathbf{P}_{t-p+1} + \left(\rho - \mathbf{I} \right) \mathbf{P}_{t-1} + \varepsilon_t \quad\ldots\ldots\ldots..28$$

But for $\rho = \left(\Phi_1 + \Phi_2 + \Phi_3 + ... + \Phi_P \right)$ then $\left(\rho - \mathbf{I} \right)$ is equal to

$$\rho - \mathbf{I} = \left(\Phi_1 + \Phi_2 + \Phi_3 + ... + \Phi_P \right) = -\left(\mathbf{I} - \Phi_1 - \Phi_2 - \Phi_3 - ... - \Phi_P \right) = -\Phi(1) \quad\ldots\ldots\ldots..29$$

Replacing equation 29 to equation 28 we will get

$$\Delta \mathbf{P}_t = \delta^* + \pi^* t + \zeta_1 \Delta \mathbf{P}_{t-1} + \zeta_2 \Delta \mathbf{P}_{t-2} + ... + \zeta_{P-1} \Delta \mathbf{P}_{t-p+1} - \Phi(1) \mathbf{P}_{t-1} + \varepsilon_t \quad\ldots\ldots\ldots..30$$

But from equation 25 $\Phi(1) = \mathbf{B}\mathbf{A'}$ is true so

$$\Delta \mathbf{P}_t = \delta^* + \pi^* t + \zeta_1 \Delta \mathbf{P}_{t-1} + \zeta_2 \Delta \mathbf{P}_{t-2} + ... + \zeta_{P-1} \Delta \mathbf{P}_{t-p+1} - \mathbf{B}\mathbf{A'} \mathbf{P}_{t-1} + \varepsilon_t \quad\ldots\ldots\ldots..31$$

This is Granger error correction representation of cointegrating VAR or Vector Error Correction Model as defined in Engle and Granger (1987) under Granger representation theorem. Assuming there are $n = h+1$ cointegrated prices, the $(n \times h)$ matrix of **A** will determine the long run relationship between $n = h+1$ prices and the $(n \times h)$ matrix of **B** will give us the adjustment parameter need to correct the prices in case any one of them deviate out of their long run relationship.

In this paper the search for cointegrating prices will start by identifying two markets prices which share the same common trend given white noise innovations. Then third, fourth and more markets are added if the new market can share the same single random trend with cointegrated markets in lower dimension. On all steps the assumption of normality and lack of serial correlation is assured by using appropriate test to be explained below. Moreover four different information criterions are used to see if the selected combination of markets has lag order backed by the information criterions. And for selected markets ARCH test is done on error terms. Once $n = h+1$ markets with h cointegrating vectors are found, Johansen method will enable us to identify the space spanned by the cointegrating vectors not the cointegrating vectors themselves. So appropriate normalization will be imposed based on proximity of the actual estimation to the ideal assumptions assumed in the model, mainly the normality of the error vectors. Even after the $n = h+1$ markets and their appropriate lag order is identified, out of P_n^n permutation of markets only $n-1$ permutations are relevant and others are just redundant. And still how these equations are ordered is found to matter for normality of the error vector.

2.4. Johansen full information Maximum likelihood estimation of vector error correction model (VECM)

The Johansen Maximum likelihood estimation can be done based on different assumption about the existence of drift and trend term in the error correction model and cointegration relationships specified in equation 31 above. In economic terms the five models are related to the assumption about the nature of the level data and the pattern of the log of transaction cost need for creating space utility. If the data does not have deterministic trend and the log of transaction costs are assumed to be zero the model with out any deterministic terms has to be used. However it is very illogical to assume that log of transaction cost to be zero in spatially dispersed markets. So this model is not relevant. The second model is related to the assumption of constant average log of transaction cost with out any deterministic trend in level data. The third model in addition to constant log of transaction cost will allow for deterministic trend in level data. The first model is called completely restricted model, the second one is called model with restricted constant and the third model is called model with unrestricted constant.

However if there is trend in log of transaction cost the use of restricted trend model will be advisable. And lastly if there is quadratic deterministic trend in the level data, which is rare in economics, the fully specified unrestricted model will be used and this is known as unrestricted trend or simply unrestricted model. As will be observed in the analysis part

there is no visible deterministic linear trend let alone a quadratic trend in the data. So the logical models are either restricted constant or restricted trend. However in order to allow for capacity to control for seasonal impacts with out allowing trend in log of transaction cost it is also advisable to use unrestricted constant model, too. Even though only 3 models are used in this study all 5 models are defined and explained below. This is so since the exposition did not result on any additional cost. The same three general estimation methods have to be defined for the 5 models and 5 of them can be extracted with some restrictions imposed on the three methods of estimation. But before going to the detail, let's first redefine the following vectors and matrixes as $-\mathbf{BA}^{'} = \zeta_0$, $\boldsymbol{\delta}* = \boldsymbol{\alpha}$ and $\boldsymbol{\pi}*=\boldsymbol{\beta}$. And the 5 models are

1. Unrestricted model or unrestricted trend model (model 1)

$$\Delta \mathbf{P}_t = \boldsymbol{\alpha} + \boldsymbol{\beta} t + \zeta_1 \Delta \mathbf{P}_{t-1} + \zeta_2 \Delta \mathbf{P}_{t-2} + ... + \zeta_{P-1} \Delta \mathbf{P}_{t-p+1} + \zeta_0 \mathbf{P}_{t-1} + \boldsymbol{\varepsilon}_t \dots \dots \dots 32$$

2. Model with out trend or model with unrestricted constant (model 2)

$$\Delta \mathbf{P}_t = \boldsymbol{\alpha} + \zeta_1 \Delta \mathbf{P}_{t-1} + \zeta_2 \Delta \mathbf{P}_{t-2} + ... + \zeta_{P-1} \Delta \mathbf{P}_{t-p+1} + \zeta_0 \mathbf{P}_{t-1} + \boldsymbol{\varepsilon}_t \dots \dots \dots 33$$

3. Model with out trend and drift or fully restricted model (model 3)

$$\Delta \mathbf{P}_t = \zeta_1 \Delta \mathbf{P}_{t-1} + \zeta_2 \Delta \mathbf{P}_{t-2} + ... + \zeta_{P-1} \Delta \mathbf{P}_{t-p+1} + \zeta_0 \mathbf{P}_{t-1} + \boldsymbol{\varepsilon}_t \dots \dots \dots 34$$

4. Model with restricted drift which is restricted to the h cointegration equations only or restricted constant model (model 4)

$$\Delta \mathbf{P}_t = \boldsymbol{\alpha}_h + \zeta_1 \Delta \mathbf{P}_{t-1} + \zeta_2 \Delta \mathbf{P}_{t-2} + ... + \zeta_{P-1} \Delta \mathbf{P}_{t-p+1} + \zeta_0 \mathbf{P}_{t-1} + \boldsymbol{\varepsilon}_t \dots \dots \dots 35$$

5. Model with restricted trend in which trend is restricted to the h cointegration equations only (model 5)

$$\Delta \mathbf{P}_t = \boldsymbol{\alpha} + \boldsymbol{\beta}_h t + \zeta_1 \Delta \mathbf{P}_{t-1} + \zeta_2 \Delta \mathbf{P}_{t-2} + ... + \zeta_{P-1} \Delta \mathbf{P}_{t-p+1} + \zeta_0 \mathbf{P}_{t-1} + \boldsymbol{\varepsilon}_t \dots \dots \dots 36$$

Since $-\mathbf{BA}^{'} = \zeta_0$ have less than full rank and the equations are having none linear parameters $\left(i.e. \ -\mathbf{BA}^{'} = \zeta_0 \right)$ Johansen reduced rank full information maximum likelihood estimation will be used. The estimation procedure is explained below.

2.4.1. <u>Johansen maximum likelihood estimation procedure for model one, two & three</u>

For model one, two and three the following general Johansen full information maximum likelihood estimation can be used. Let's first define the following OLS regressions

$$\Delta \mathbf{P}_t = \boldsymbol{\theta} + \boldsymbol{\eta} \ t + \xi_1 \Delta \mathbf{P}_{t-1} + \xi_2 \Delta \mathbf{P}_{t-2} + ... + \xi_{P-1} \Delta \mathbf{P}_{t-p+1} + \mathbf{u}_t \dots \dots \dots 37$$

$$\mathbf{P}_{t-1} = \varphi + \kappa\, \mathbf{t} + \chi_1 \Delta\mathbf{P}_{t-1} + \chi_2 \Delta\mathbf{P}_{t-2} + ... + \chi_{P-1}\Delta\mathbf{P}_{t-p+1} + \mathbf{V}_t \quad\text{...........................38}$$
So
$$\alpha = \theta - \zeta_0 \varphi \quad\text{...39}$$
$$\beta = \eta - \zeta_0 \kappa \quad\text{...40}$$
$$\zeta_i = \xi_i - \zeta_0 \chi_i \quad for \quad i = 0,1,2,....P-1 \quad\text{.........................41}$$

Means for given value of ζ_0, the drift (α), trend coefficient (β) and coefficients of the auto regressive variables (ζ_i) can be derived from the above OLS regression coefficients. But still the value of ζ_0 which can give us the true value for the above coefficients have to be estimated some how. This can be done by using the fully concentrated log likelihood function based on canonical correlations. For $\sum_{XY} = E(XY)$, let's define the following matrixes

$$\left(\sum_{\mathbf{v}_t\mathbf{v}_t}\right)^{-1}\left(\sum_{\mathbf{v}_t\mathbf{u}_t}\right)\left(\sum_{\mathbf{u}_t\mathbf{u}_t}\right)^{-1}\left(\sum_{\mathbf{u}_t\mathbf{v}_t}\right)\quad\text{.......................................42}$$

$$\left(\sum_{\mathbf{u}_t\mathbf{u}_t}\right)^{-1}\left(\sum_{\mathbf{u}_t\mathbf{v}_t}\right)\left(\sum_{\mathbf{v}_t\mathbf{v}_t}\right)^{-1}\left(\sum_{\mathbf{v}_t\mathbf{u}_t}\right)\quad\text{.......................................43}$$

The Eigen values of the two matrixes given above are the same and equal to eigen vector λ but the related characteristics vectors are different. For Eigen values of equation 42 let's define the associated Eigen vectors as $\mathbf{m'}_i$. If the characteristic vectors of equation 42 are collected in descending order of their Eigen value under matrix $\mathbf{M'}$, then $\mathbf{M'}$ will be equal to $\left[\mathbf{m_1'},\mathbf{m_2'},\mathbf{m_3'}......\mathbf{m_n'}\right]$. But like any characteristics vectors the space spanned by these vectors, but not the vectors them selves can be identified. So the following normalization condition will be imposed on them. The normalization condition is given by $\mathbf{M'}\sum_{v_t v_t}\mathbf{M} = \mathbf{I}$.

And the related characteristics vectors for equation 43 are given by $\mathbf{h'}_i$ as identified by $\mathbf{H'}\sum_{\mathbf{u}_t\mathbf{u}_t}\mathbf{H} = \mathbf{I}$ after being ordered in descending order by their Eigen values. Means the matrix $\mathbf{H}$ is defined as $\mathbf{H'} = \left[\mathbf{h_1'},\mathbf{h_2'},\mathbf{h_3'}......\mathbf{h_n'}\right]$. The error term in equation 37 or $\mathbf{u}_t$ is and the error term in equation 38 or $\mathbf{V}_t$ are n dimensional vectors. Let's define

$$\mathbf{H'u}_t = \boldsymbol{\rho}_t \text{ and } \mathbf{M'V}_t = \boldsymbol{\tau}_t \quad\text{.......................................44}$$
As result $\boldsymbol{\rho}_t$ and $\boldsymbol{\tau}_t$ are n dimensional vectors each, in which

$$E(\boldsymbol{\rho}_t\boldsymbol{\rho}_t) = \mathbf{I}, \quad E(\boldsymbol{\tau}_t\boldsymbol{\tau}_t) = \mathbf{I}, \quad E(\boldsymbol{\rho}_t\boldsymbol{\tau}_t) = \mathbf{R} \quad\text{.......................45}$$

Where $\mathbf{R}$ is $(n\times n)$ diagonal matrix, in which the diagonal elements are equal to the square root of the Eigen values of equation 42 or 43, when ordered in descending order. Formally

$$
\mathbf{R} = \begin{Bmatrix} r_1 & 0 & 0 & . & . & . & 0 & 0 \\ 0 & r_2 & 0 & . & . & . & 0 & 0 \\ 0 & 0 & r_3 & . & . & . & 0 & 0 \\ . & & & & & & & . \\ . & & & & & & & . \\ . & & & & & & & . \\ 0 & 0 & 0 & . & . & . & r_{n-1} & 0 \\ 0 & 0 & 0 & . & . & . & . & r_n \end{Bmatrix} = \begin{Bmatrix} \sqrt{\lambda_1} & 0 & 0 & . & . & . & 0 & 0 \\ 0 & \sqrt{\lambda_2} & 0 & . & . & . & 0 & 0 \\ 0 & 0 & \sqrt{\lambda_3} & . & . & . & 0 & 0 \\ . & & & & & & & . \\ . & & & & & & & . \\ . & & & & & & & . \\ 0 & 0 & 0 & . & . & . & \sqrt{\lambda_{n-1}} & 0 \\ 0 & 0 & 0 & . & . & . & . & \sqrt{\lambda_n} \end{Bmatrix} \quad \ldots\ldots 46
$$

Now let's replicate the unrestricted error correction model of equation 32 below

$$
\Delta \mathbf{P}_t = \boldsymbol{\alpha} + \boldsymbol{\beta} t + \boldsymbol{\zeta}_1 \Delta \mathbf{P}_{t-1} + \boldsymbol{\zeta}_2 \Delta \mathbf{P}_{t-2} + \ldots + \boldsymbol{\zeta}_{P-1} \Delta \mathbf{P}_{t-p+1} + \boldsymbol{\zeta}_0 \mathbf{P}_{t-1} + \boldsymbol{\varepsilon}_t \ldots\ldots\ldots\ldots .32
$$

And the error term $\left(\boldsymbol{\varepsilon}_t \right)$ will be equal to

$$
\boldsymbol{\varepsilon}_t = \Delta \mathbf{P}_t - \boldsymbol{\alpha} - \boldsymbol{\beta} t - \boldsymbol{\zeta}_1 \Delta \mathbf{P}_{t-1} - \boldsymbol{\zeta}_2 \Delta \mathbf{P}_{t-2} - \ldots - \boldsymbol{\zeta}_{P-1} \Delta \mathbf{P}_{t-p+1} - \boldsymbol{\zeta}_0 \mathbf{P}_{t-1} = \mathbf{u}_t - \boldsymbol{\zeta}_0 \mathbf{v}_t \ldots\ldots .47
$$

Let's focus on the full information maximum likelihood function given by

$$
\mathbf{LL} = -\frac{\mathbf{Tn}}{2} \log(2\pi) - \frac{\mathbf{T}}{2} \log|\boldsymbol{\Omega}| - \frac{\mathbf{T}}{2} \sum_{t=1}^{\mathrm{T}} \boldsymbol{\varepsilon}_t' \boldsymbol{\Omega}^{-1} \boldsymbol{\varepsilon}_t \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots .48
$$

Given $\boldsymbol{\Omega}$ is variance and covariance matrix of equation 32 and maximization of the above log likelihood function will result on estimation of all the needed coefficients. But the problem is that the adjustment parameters $\left(\mathbf{B} \right)$ and cointegration coefficients $\left(\mathbf{A} \right)$ are non linear function of the vector of error terms $\boldsymbol{\varepsilon}_t$. And the estimation of the coefficients will demand much complicated algorithm of reduced rank regression in none liner form. However observing the fact maximization equation 48 is the same as maximization of the following concentrated log likelihood function.

$$
\mathbf{LL}(\boldsymbol{\zeta}_0) = -\frac{\mathbf{Tn}}{2} \log(2\pi) - \frac{\mathbf{T}}{2} \log|\boldsymbol{\Omega}| - \frac{\mathbf{T}}{2} \sum_{t=1}^{\mathrm{T}} \left(\mathbf{u}_t - \boldsymbol{\zeta}_0 \mathbf{v}_t \right)' \boldsymbol{\Omega}^{-1} \left(\mathbf{u}_t - \boldsymbol{\zeta}_0 \mathbf{v}_t \right) \ldots\ldots\ldots\ldots .49
$$

The variance covariance matrix $\left(\boldsymbol{\Omega} \right)$ which maximize the above log likelihood for given ζ_0 is equal to

$$
\boldsymbol{\Omega}(\boldsymbol{\zeta}_0) = \frac{1}{\mathbf{T}} \sum_{t=1}^{\mathrm{T}} \left(\mathbf{u}_t - \boldsymbol{\zeta}_0 \mathbf{v}_t \right) \left(\mathbf{u}_t - \boldsymbol{\zeta}_0 \mathbf{v}_t \right)' \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots .50
$$

And if equation 50 is replaced in to the right last part in equation 49, then it will give us

$$
\frac{\mathbf{T}}{2} \sum_{t=1}^{\mathrm{T}} \left(\mathbf{u}_t - \boldsymbol{\zeta}_0 \mathbf{v}_t \right)' \left(\frac{1}{\mathbf{T}} \sum_{t=1}^{\mathrm{T}} \left(\mathbf{u}_t - \boldsymbol{\zeta}_0 \mathbf{v}_t \right) \left(\mathbf{u}_t - \boldsymbol{\zeta}_0 \mathbf{v}_t \right)' \right)^{-1} \left(\mathbf{u}_t - \boldsymbol{\zeta}_0 \mathbf{v}_t \right) = \frac{\mathbf{Tn}}{2} \ldots\ldots\ldots\ldots .51
$$

Using equation 49, 50 and 51 the fully concentrated version of the maximum likelihood

equation in 48 for given value of ζ_0 and Ω will be

$$\mathbf{LL}(\zeta_0,\Omega)=-\frac{\mathbf{Tn}}{2}\log(2\pi)-\frac{\mathbf{T}}{2}\log\left|\frac{1}{\mathbf{T}}\sum_{t=1}^{\mathbf{T}}(\mathbf{u}_t-\zeta_0\mathbf{v}_t)(\mathbf{u}_t-\zeta_0\mathbf{v}_t)'\right|-\frac{\mathbf{Tn}}{2}$$

$$\mathbf{LL}(\zeta_0,\Omega)=-\frac{\mathbf{Tn}}{2}\log(2\pi)-\frac{\mathbf{Tn}}{2}-\frac{\mathbf{T}}{2}\log\left|\frac{1}{\mathbf{T}}\sum_{t=1}^{\mathbf{T}}(\mathbf{u}_t-\zeta_0\mathbf{v}_t)(\mathbf{u}_t-\zeta_0\mathbf{v}_t)'\right| \quad\ldots\ldots\ldots\ldots\text{52}$$

But the OLS error terms can be written as $\mathbf{u}_t=(\mathbf{H'})^{-1}\boldsymbol{\rho}_t$ and $\mathbf{v}_t=(\mathbf{M'})^{-1}\boldsymbol{\tau}_t$ based on equation 44. Given the new information the right last part of equation 52 will be

$$\frac{\mathbf{T}}{2}\log\left|\frac{1}{\mathbf{T}}\sum_{t=1}^{\mathbf{T}}\left((\mathbf{H'})^{-1}\boldsymbol{\rho}_t-\zeta_0(\mathbf{M'})^{-1}\boldsymbol{\tau_t}\right)\left((\mathbf{H'})^{-1}\boldsymbol{\rho}_t-\zeta_0(\mathbf{M'})^{-1}\boldsymbol{\tau_t}\right)'\right|$$

$$\frac{\mathbf{T}}{2}\log\left|(\mathbf{H'})^{-1}\frac{1}{\mathbf{T}}\sum_{t=1}^{\mathbf{T}}\left(\boldsymbol{\rho}_t-(\mathbf{H'})\zeta_0(\mathbf{M'})^{-1}\boldsymbol{\tau_t}\right)\left(\boldsymbol{\rho}_t-(\mathbf{H'})\zeta_0(\mathbf{M'})^{-1}\boldsymbol{\tau_t}\right)'(\mathbf{H'})^{-1}\right|$$

$$\frac{\mathbf{T}}{2}\log\left|\frac{1}{\mathbf{T}}\sum_{t=1}^{\mathbf{T}}\left(\boldsymbol{\rho}_t-(\mathbf{H'})\zeta_0(\mathbf{M'})^{-1}\boldsymbol{\tau_t}\right)\left(\boldsymbol{\rho}_t-(\mathbf{H'})\zeta_0(\mathbf{M'})^{-1}\boldsymbol{\tau_t}\right)'\right|\left|(\mathbf{H'})^{-1}\right|^2$$

$$\frac{\mathbf{T}}{2}\log\left|\frac{1}{\mathbf{T}}\sum_{t=1}^{\mathbf{T}}(\boldsymbol{\rho}_t-\mathbf{\Pi}\boldsymbol{\tau_t})(\boldsymbol{\rho}_t-\mathbf{\Pi}\boldsymbol{\tau_t})'\right|\left|(\mathbf{H'})^{-1}\right|^2 \quad\ldots\ldots\ldots\ldots\ldots\ldots\ldots\text{53}$$

Given $(\mathbf{H'})\zeta_0(\mathbf{M'})^{-1}=\mathbf{\Pi}$. But for single price (i) in period t the expression $\frac{1}{T}\sum_{t=1}^{T}(\boldsymbol{\rho}_{it}-\Pi\tau_{it})(\boldsymbol{\rho}_{it}-\Pi\tau_{it})'$ is average sum of square of the residual of regression of ρ_{it} on τ_{it}. And this is equal to $1-r_i^2$ for r_i^2 is degree of determination related with price i. And $1-r_i^2=1-\lambda_i$ from equation 46. Moreover given from equation 45 $\boldsymbol{\rho}_t$ and $\boldsymbol{\tau}_\tau$ are orthogonal for $t\neq\tau$, and there are $h=n-1$ cointegrating vectors, the following fact will naturally follow

$$\left|\frac{1}{T}\sum_{t=1}^{T}(\boldsymbol{\rho}_t-\mathbf{\Pi}\boldsymbol{\tau}_t)(\boldsymbol{\rho}_t-\mathbf{\Pi}\boldsymbol{\tau}_t)'\right|=\begin{vmatrix}1-\lambda_1 & 0 & 0 & . & . & . & 0 & 0\\ 0 & 1-\lambda_2 & 0 & . & . & . & 0 & 0\\ 0 & 0 & 1-\lambda_3 & . & . & . & 0 & 0\\ . & . & . & . & . & . & . & .\\ . & . & . & . & . & . & . & .\\ . & . & . & . & . & . & . & .\\ 0 & 0 & 0 & . & . & . & 1-\lambda_h & 0\\ 0 & 0 & 0 & . & . & . & . & 1-0\end{vmatrix}=\prod_{i=1}^{h}(1-\lambda_i)$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\text{54}$$

Inserting equation 54 in to equation 53results on

$$\frac{T}{2}\log\prod_{i=1}^{h}(1-\lambda_i)\div\left|\left(\mathbf{H}^{'}\right)\right|^{2}$$ 55

Additionally since to identify the characteristics vectors of equation 43 the normalization of $\mathbf{H}^{'}\sum_{\mathbf{u_t u_t}}\mathbf{H}=\mathbf{I}$ is imposed. So using this fact the last part f equation 55 will be

$$\left|\mathbf{H}^{'}\sum_{\mathbf{u_t u_t}}\mathbf{H}\right|=\left|\mathbf{I}\right|=1$$

$$\left|\mathbf{H}^{'}\right|\left|\sum_{\mathbf{u_t u_t}}\right|\left|\mathbf{H}\right|=1$$

$$\left|\sum_{\mathbf{u_t u_t}}\right|\left|\mathbf{H}\right|^{2}=1$$

$$\left|\sum_{\mathbf{u_t u_t}}\right|=1/\left|\mathbf{H}\right|^{2}$$ 56

Inserting equation 56 in to equation 55 will result on

$$\frac{T}{2}\log\prod_{i=1}^{h}(1-\lambda_i)\left|\sum_{\mathbf{u_t u_t}}\right|$$

$$\frac{T}{2}\log\prod_{i=1}^{h}(1-\lambda_i)+\frac{T}{2}\log\left|\sum_{\mathbf{u_t u_t}}\right|$$ 57

By replacing equation 57 in to equation 52 will give us the fully concentrated full information Johansen maximum log likelihood for unrestricted model or model 1 given in equation 32

$$LL=-\frac{Tn}{2}\log(2\pi)-\frac{Tn}{2}-\frac{T}{2}\log\left|\sum_{\mathbf{u_t u_t}}\right|-\frac{T}{2}\log\prod_{i=1}^{h}(1-\lambda_i)$$ 58

But how do we get the adjustment parameter and the long run coefficients from the above maximum likelihood function, in addition to other coefficients in the vector error correction model? The space spanned by h cointegration equations in $\mathbf{A}$ is given by h Eigen vectors associated with h largest Eigen values of the matrix in equation 42. Or the first h vectors of $\mathbf{M}^{'}$. However what we are estimating is a liner combination of the cointegration vectors of the same dimension or space spanned by cointegration vectors not the cointegration vectors themselves. So some form of normalization is needed in order to identify them. And the identifying restriction as stated above is given by $\mathbf{M}^{'}\sum_{v_t v_t}\mathbf{M}=\mathbf{I}$.

To estimate the adjustment parameters $\left(\mathbf{B}\right)$ and ζ_0 for $-\mathbf{BA}^{'}=\zeta_0$, we have to notice that the value of $\mathbf{B}$ is found by regressing $\mathbf{u}_t$ on $\mathbf{A}^{'}\mathbf{v}_t$. So for $\mathbf{A}^{'}\mathbf{v}_t=\mathbf{z}_t$ we can find the value of $\mathbf{B}$ by linear regression of $\mathbf{u}_t$ on $\mathbf{z}_t$. Formally

$$B = -\frac{\frac{1}{T}\sum_{t=1}^{T}\mathbf{u}_t \mathbf{z}_t^{'}}{\frac{1}{T}\sum_{t=1}^{T}\mathbf{z}_t \mathbf{z}_t^{'}} = -\frac{\frac{1}{T}\sum_{t=1}^{T}\mathbf{u}_t \left(A^{'}\mathbf{v}_t\right)^{'}}{I} = -\frac{1}{T}\sum_{t=1}^{T}\mathbf{u}_t \mathbf{v}_t^{'} A = -\sum_{\mathbf{u}_t \mathbf{v}_t} A \quad \text{.................59}$$

And since $\zeta_0 = -BA^{'}$

$$\zeta_0 = \sum_{\mathbf{u}_t \mathbf{v}_t} AA^{'} \quad \text{..60}$$

If the adjustment matrix $(\mathbf{B})$, the cointegration matrix $(\mathbf{A})$ and the error correction coefficient on unit root terms (ζ_0) are estimated, we can use equation 39 up to 41 to estimate the remaining coefficients.

For model 2 the coefficient of the trend term $(\boldsymbol{\beta})$ will be fixed at zero and the trend term will be dropped from equation 37, 38 and 47. And the rest is the same. For model three the coefficient of the trend term $(\boldsymbol{\beta})$ and the coefficient of drift term $(\boldsymbol{\alpha})$ will be fixed at zero in equation 37, 38 and 47. This means dropping the drift and the time variable from the three equations and the rest is the same. This is why theoretical development for process of estimation for the 5 models is as costly as developing for 3 models used in this paper.

2.4.2. <u>Johansen maximum likelihood estimation procedure for model four</u>

Unlike for model 2 and 3, for the restricted drift and restricted trend model some modification is needed in estimation process. In restricted drift model by assumption drift is found in the cointegration equations but not in the error correction model. To explain how estimation is done under such restrictions let's replicated equation 35, below.

$$\Delta\mathbf{P}_t = \boldsymbol{\alpha}_h + \zeta_1 \Delta\mathbf{P}_{t-1} + \zeta_2 \Delta\mathbf{P}_{t-2} + ... + \zeta_{P-1}\Delta\mathbf{P}_{t-p+1} + \zeta_0 \mathbf{P}_{t-1} + \boldsymbol{\varepsilon}_t \quad \text{..................35}$$

Assuming the cointegration vector $(\mathbf{Z})$ is equal to $\mathbf{Z} = \boldsymbol{\kappa} - A^{'}\mathbf{P}_{T-1}$, it will follow that $\boldsymbol{\alpha}_h = \mathbf{B}\boldsymbol{\kappa}$. Given the above fact and $-BA^{'} = \zeta_0$ in equation 35 can be re-parameterized as

$$\Delta\mathbf{P}_t = \mathbf{B}\boldsymbol{\kappa} + \zeta_1 \Delta\mathbf{P}_{t-1} + \zeta_2 \Delta\mathbf{P}_{t-2} + ... + \zeta_{P-1}\Delta\mathbf{P}_{t-p+1} - BA^{'}\mathbf{P}_{t-1} + \boldsymbol{\varepsilon}_t$$

$$\Delta\mathbf{P}_t = \zeta_1 \Delta\mathbf{P}_{t-1} + \zeta_2 \Delta\mathbf{P}_{t-2} + ... + \zeta_{P-1}\Delta\mathbf{P}_{t-p+1} + B\left(\boldsymbol{\kappa} - A^{'}\mathbf{P}_{t-1}\right) + \boldsymbol{\varepsilon}_t \quad \text{...............61}$$

To find the parameters of equation 61, we can use the following OLS residuals

$$\Delta\mathbf{P}_t = \xi_1 \Delta\mathbf{P}_{t-1} + \xi_2 \Delta\mathbf{P}_{t-2} + ... + \xi_{P-1}\Delta\mathbf{P}_{t-p+1} + \mathbf{u}_t \quad \text{.............................62}$$

$$\mathbf{P}_{t-1} = \chi_1 \Delta\mathbf{P}_{t-1} + \chi_2 \Delta\mathbf{P}_{t-2} + ... + \chi_{P-1}\Delta\mathbf{P}_{t-p+1} + \mathbf{v}_t \quad \text{............................63}$$

$$i = \varpi_1 \Delta \mathbf{P}_{t-1} + \varpi_2 \Delta \mathbf{P}_{t-2} + ... + \varpi_{P-1} \Delta \mathbf{P}_{t-p+1} + \mathbf{W}_t \text{...}64$$

Given i is n dimensional vector of ones. For the vector of error terms (ε_t) in equation 61 and related variance covariance matrix $(\mathbf{\Omega})$, based on equations numbered from 62 to 64, the following facts are true.

$$\varepsilon_t = \mathbf{u}_t - \zeta_0 \mathbf{V}_t - \alpha_h \mathbf{W}_t$$

$$\mathbf{\Omega}(\zeta_0, \alpha_h) = \frac{1}{T} \sum_{t=1}^{T} (\varepsilon_t)(\varepsilon_t)' = \frac{1}{T} \sum_{t=1}^{T} (\mathbf{u}_t - \zeta_0 \mathbf{V}_t - \alpha_h \mathbf{W}_t)(\mathbf{u}_t - \zeta_0 \mathbf{V}_t - \alpha_h \mathbf{W}_t)'$$

So the fully concentrated maximum log likelihood function is give as

$$LL = -\frac{Tn}{2}\log(2\pi) - \frac{Tn}{2} - \frac{T}{2}\log\left|\sum_{t=1}^{T}(\mathbf{u}_t - \zeta_0 \mathbf{V}_t - \alpha_h \mathbf{W}_t)(\mathbf{u}_t - \zeta_0 \mathbf{V}_t - \alpha_h \mathbf{W}_t)'\right|\text{......}65$$

But in equation 61 it is assumed that $\alpha_h = \mathbf{B}\kappa$ and $-\mathbf{BA}' = \zeta_0$. So the matrix $\mathbf{u}_t - \zeta_0 \mathbf{V}_t - \alpha_h \mathbf{W}_t$ can be expressed as

$$\mathbf{u}_t - \zeta_0 \mathbf{V}_t - \alpha_h \mathbf{W}_t = \mathbf{u}_t - \mathbf{B}\kappa \mathbf{W}_t + \mathbf{BA}'v_t = \mathbf{u}_t + \mathbf{B}\begin{vmatrix} -\kappa & \mathbf{A}' \end{vmatrix}\begin{vmatrix} \mathbf{W}_t \\ \mathbf{V}_t \end{vmatrix} = \mathbf{u}_t + \mathbf{B}\hat{\mathbf{A}}'\mathcal{W}_t.$$

Inserting the above result in to equation 65 will give us

$$LL = -\frac{Tn}{2}\log(2\pi) - \frac{Tn}{2} - \frac{T}{2}\log\left|\frac{1}{T}\sum_{t=1}^{T}\left(\mathbf{u}_t + \mathbf{B}\hat{\mathbf{A}}'\mathcal{W}\right)\left(\mathbf{u}_t + \mathbf{B}\hat{\mathbf{A}}'\mathcal{w}\right)'\right|\text{..............}66$$

For $\mathcal{W}_t' = \begin{bmatrix} \mathbf{w}_t & \mathbf{V}_t \end{bmatrix}$ in which $\mathcal{W}_t$ is $n+1$ dimension vector and $\mathbf{u}_t$ is an $n+1$ dimension vector. We can use two matrixes given in equation 67 and equation 68, below, to find the basses of the cointegrating vectors and the largest Eigen values which can be used to maximize equation 66.

$$\left(\sum_{\mathbf{w}_t\mathbf{w}_t}\right)^{-1}\left(\sum_{\mathbf{w}_t\mathbf{u}_t}\right)\left(\sum_{\mathbf{u}_t\mathbf{u}_t}\right)^{-1}\left(\sum_{\mathbf{u}_t\mathbf{w}_t}\right)\text{..}67$$

$$\left(\sum_{\mathbf{u}_t\mathbf{u}_t}\right)^{-1}\left(\sum_{\mathbf{u}_t\mathbf{w}_t}\right)\left(\sum_{\mathbf{w}_t\mathbf{w}_t}\right)^{-1}\left(\sum_{\mathbf{w}_t\mathbf{u}_t}\right)\text{..}68$$

The Eigen vectors related to the Eigen values of the matrix in equation 67 ordered in descending order will give us the basses of the $\mathbf{m}_i$ Eigen vectors for $i = 1, 2, 3 ... n+1$. And define $\mathbf{M}' = \{\mathbf{m}_1', \mathbf{m}_2', \mathbf{m}_3'.....\mathbf{m}_{n+1}'\}$ normalized by $\mathbf{M}'\left(\sum_{\mathbf{w}_t\mathbf{w}_t}\right)\mathbf{M} = \mathbf{I}$. And again using equation 68 we can drive the basses of the Eigen vectors, ordered in descending order of their Eigen value. And if they are stacked in $\mathbf{H}'$ as $\mathbf{H}' = \{\mathbf{h}_1', \mathbf{h}_2', \mathbf{h}_3'.....\mathbf{h}_n'\}$ and are normalized by $\mathbf{H}'\left(\sum_{\mathbf{u}_t\mathbf{u}_t}\right)\mathbf{H} = \mathbf{I}$ will give us the matrix of Eigen vectors $\mathbf{H}'$ related

to matrix in equation 68. Then by definition of canonical correlation for $\mathbf{H}'\mathbf{u_t} = \boldsymbol{\rho}_t$ and $\mathbf{M}'\mathbf{w_t} = \boldsymbol{\tau}_t$ it follows that

$$E\left(\boldsymbol{\rho}_t\boldsymbol{\rho}_t\right) = \mathbf{I} \; , \; E\left(\boldsymbol{\tau}_t\boldsymbol{\tau}_t\right) = \mathbf{I} \; \text{and} \; E\left(\boldsymbol{\rho}_t\boldsymbol{\tau}_t\right) = \mathbf{R} \; \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots 69$$

Following the same logic as before the fully concentrated full information maximum likelihood would be

$$LL = -\frac{Tn}{2}\log(2\pi) - \frac{Tn}{2} - \frac{T}{2}\log\left|\Sigma_{\mathbf{u_t u_t}}\right| - \frac{T}{2}\log\prod_{i=1}^{h}(1-\lambda_i) \; \dots\dots\dots\dots\dots 70$$

This is the same as equation 58 except for the difference in canonical terms introduced in equation 69 related to equation 44. So the coefficients of equation 35 or equation 61 are found using the coefficients in equation 62 to 64 for given value of ζ_0, as

$$\zeta_i = \xi_i - \zeta_0\chi_i - \boldsymbol{\alpha}\,\varpi_i \qquad \text{for } i = 0,1,2\dots \dots P-1 \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots 71$$

The cointegration vector or $\widehat{\mathbf{A}}$ is the collection of the first h Eigen vectors associated with $\mathbf{M}'$ or $\widehat{\mathbf{A}}' = \left\{\mathbf{m}_1',\mathbf{m}_2',\mathbf{m}_3'\dots\mathbf{m}_h'\right\}$ when they are ordered in descending order by size of their Eigen value. Following the same logic as equation 59, the adjustment parameter $\mathbf{B}$ is given by equation 72 below

$$\mathbf{B} = -\Sigma_{\mathbf{u_t w_t}}\,\widehat{\mathbf{A}} = \Sigma_{\mathbf{u_t w_t}}\begin{bmatrix}\boldsymbol{\kappa} \\ \mathbf{-A}\end{bmatrix} \; \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots 72$$

And the rest of the coefficients are found by noticing that $\mathbf{B}\widehat{\mathbf{A}}'$ is composed of intercept $\left(\boldsymbol{\alpha}\right)$ and the coefficient of the cointegration equation $\left(\zeta_0\right)$

$$\mathbf{B}\widehat{\mathbf{A}}' = -\Sigma_{\mathbf{u_t w_t}}\,\widehat{\mathbf{A}}\widehat{\mathbf{A}}' = \mathbf{B}\begin{bmatrix}-\boldsymbol{\kappa} & \mathbf{A}'\end{bmatrix} = \begin{bmatrix}-\mathbf{B}\boldsymbol{\kappa} & \mathbf{B}\mathbf{A}'\end{bmatrix} = \begin{bmatrix}-\boldsymbol{\alpha} & -\zeta_0\end{bmatrix}\dots\dots\dots\dots\dots 73$$

And the variance covariance matrix as stated above is give by

$$\boldsymbol{\Omega} = \frac{1}{T}\sum_{t=1}^{T}(\boldsymbol{\varepsilon}_t)(\boldsymbol{\varepsilon}_t)' = \frac{1}{T}\sum_{t=1}^{T}(\mathbf{u}_t - \zeta_0\mathbf{v}_t - \boldsymbol{\alpha}_h\mathbf{W}_t)(\mathbf{u}_t - \zeta_0\mathbf{v}_t - \boldsymbol{\alpha}_h\mathbf{W}_t)' \; \dots\dots\dots\dots\dots 74$$

So the above procedures will give us the estimate of the coefficients of the restricted drift model specified in model 4 or in equation 35.

2.4.3. Johansen maximum likelihood estimation procedure for model five

And the last model of restricted trend given in equation 36 and is replicated below

$$\Delta\mathbf{P}_t = \boldsymbol{\alpha} + \boldsymbol{\beta}_h\mathbf{t} + \zeta_1\Delta\mathbf{P}_{t-1} + \zeta_2\Delta\mathbf{P}_{t-2} + \dots + \zeta_{P-1}\Delta\mathbf{P}_{t-p+1} + \zeta_0\mathbf{P}_{t-1} + \boldsymbol{\varepsilon}_t \; \dots\dots\dots\dots\dots 36$$

For $\boldsymbol{\alpha} = \boldsymbol{\alpha}_0 + \boldsymbol{\alpha}_h = \boldsymbol{\alpha}_0 + \mathbf{B}\boldsymbol{\kappa}$, $\boldsymbol{\beta}_h = \mathbf{B}\boldsymbol{\psi}$ and $\boldsymbol{\zeta}_0 = -\mathbf{B}\mathbf{A}'$ equation 36 given new definition of the cointegration equations as $\mathbf{Z} = \boldsymbol{\kappa} + \boldsymbol{\psi}\mathbf{t} - \mathbf{A}'\mathbf{P}_{t-1}$ can be re-parameterize as

$$\Delta\mathbf{P}_t = \boldsymbol{\alpha}_0 + \boldsymbol{\zeta}_1\Delta\mathbf{P}_{t-1} + \boldsymbol{\zeta}_2\Delta\mathbf{P}_{t-2} + ... + \boldsymbol{\zeta}_{P-1}\Delta\mathbf{P}_{t-p+1} + \mathbf{B}\boldsymbol{\kappa} + \mathbf{B}\boldsymbol{\psi}t - \mathbf{B}\mathbf{A}'\mathbf{P}_{t-1} + \boldsymbol{\varepsilon}_t$$

$$\Delta\mathbf{P}_t = \boldsymbol{\alpha}_0 + \boldsymbol{\zeta}_1\Delta\mathbf{P}_{t-1} + \boldsymbol{\zeta}_2\Delta\mathbf{P}_{t-2} + ... + \boldsymbol{\zeta}_{P-1}\Delta\mathbf{P}_{t-p+1} + \mathbf{B}\left(\boldsymbol{\kappa} + \boldsymbol{\psi}t - \mathbf{A}'\mathbf{P}_{t-1}\right) + \boldsymbol{\varepsilon}_t \dots\dots\dots 75$$

For cointegration equations given by $\boldsymbol{\kappa} + \boldsymbol{\psi}\mathbf{t} - \mathbf{A}'\mathbf{P}_{t-1}$, the adjustment parameter is given by $\mathbf{B}$ as before. By maximizing the log likelihood in equation 76 below, we can estimate the coefficients in equation 36 or 75.

$$LL = -\frac{Tn}{2}\log(2\pi) - \frac{T}{2}\log|\boldsymbol{\Omega}| - \frac{T}{2}\sum_{t=1}^{T}\boldsymbol{\varepsilon}_t'\boldsymbol{\Omega}^{-1}\boldsymbol{\varepsilon}_t \dots\dots\dots\dots\dots\dots\dots\dots 76$$

Unfortunately equation 75 and 76 are non liner in parameters in addition to reduced rank nature of the level of prices which demands reduced rank non linear regression. But this can be avoided by using fully concentrated maximum likelihood procedure as proposed by Johansen (1988 and 1991). In order to use the Johansen fully concentrated full information maximum likelihood estimation which is liner in parameters, let's develop the following OLS regressions

$$\Delta\mathbf{P}_t = \boldsymbol{\theta} + \boldsymbol{\xi}_1\Delta\mathbf{P}_{t-1} + \boldsymbol{\xi}_2\Delta\mathbf{P}_{t-2} + ... + \boldsymbol{\xi}_{P-1}\Delta\mathbf{P}_{t-p+1} + \mathbf{u}_t \dots\dots\dots\dots\dots 77$$

$$\mathbf{P}_{t-1} = \boldsymbol{\varphi} + \boldsymbol{\chi}_1\Delta\mathbf{P}_{t-1} + \boldsymbol{\chi}_2\Delta\mathbf{P}_{t-2} + ... + \boldsymbol{\chi}_{P-1}\Delta\mathbf{P}_{t-p+1} + \mathbf{v}_t \dots\dots\dots\dots 78$$

$$\mathbf{i} = \boldsymbol{\varpi}_1\Delta\mathbf{P}_{t-1} + \boldsymbol{\varpi}_2\Delta\mathbf{P}_{t-2} + ... + \boldsymbol{\varpi}_{P-1}\Delta\mathbf{P}_{t-p+1} + \mathbf{W}_t \dots\dots\dots\dots\dots\dots\dots 79$$

$$\mathbf{t} = \boldsymbol{\gamma} + \ell_1\Delta\mathbf{P}_{t-1} + \ell_2\Delta\mathbf{P}_{t-2} + ... + \ell_{P-1}\Delta\mathbf{P}_{t-p+1} + \vartheta_t \dots\dots\dots\dots\dots 80$$

And for $\boldsymbol{\varepsilon}_t = \mathbf{u}_t - \boldsymbol{\alpha}_h\mathbf{W}_t - \boldsymbol{\beta}_h\vartheta_t - \boldsymbol{\zeta}_0\mathbf{v}_t$ & for the concentrated variance covariance matrix given by $\boldsymbol{\Omega}(\boldsymbol{\alpha}_h, \boldsymbol{\beta}_h, \boldsymbol{\zeta}_0) = \frac{1}{T}\sum_{t=1}^{T}\left(\mathbf{u}_t - \boldsymbol{\alpha}_h\mathbf{W}_t - \boldsymbol{\beta}_h\vartheta_t - \boldsymbol{\zeta}_0\mathbf{v}_t\right)\left(\mathbf{u}_t - \boldsymbol{\alpha}_h\mathbf{W}_t - \boldsymbol{\beta}_h\vartheta_t - \boldsymbol{\zeta}_0\mathbf{v}_t\right)'$ the fully concentrated log likelihood function related to equation 76 will be

$$LL = -\frac{Tn}{2}\log(2\pi) - \frac{Tn}{2} - \frac{T}{2}\log\left|\frac{1}{T}\sum_{t=1}^{T}\left(\mathbf{u}_t - \boldsymbol{\alpha}_h\mathbf{W}_t - \boldsymbol{\beta}_h\vartheta_t - \boldsymbol{\zeta}_0\mathbf{v}_t\right)\left(\mathbf{u}_t - \boldsymbol{\alpha}_h\mathbf{W}_t - \boldsymbol{\beta}_h\vartheta_t - \boldsymbol{\zeta}_0\mathbf{v}_t\right)'\right|$$

$$\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots 81$$

For $\boldsymbol{\alpha}_h = \mathbf{B}\boldsymbol{\kappa}$, $\boldsymbol{\beta}_h = \mathbf{B}\boldsymbol{\psi}$ and $\boldsymbol{\zeta}_0 = -\mathbf{B}\mathbf{A}'$, the error term in equation 81 or $\mathbf{u}_t - \boldsymbol{\alpha}_h\mathbf{W}_t - \boldsymbol{\beta}_h\vartheta_t - \boldsymbol{\zeta}_0\mathbf{v}_t$ can re-parameterized as

$$\mathbf{u}_t - \boldsymbol{\alpha}_h\mathbf{W}_t - \boldsymbol{\beta}_h\vartheta_t - \boldsymbol{\zeta}_0\mathbf{v}_t = \mathbf{u}_t - \mathbf{B}\boldsymbol{\kappa}\mathbf{W}_t - \mathbf{B}\boldsymbol{\psi}\vartheta_t + \mathbf{B}\mathbf{A}'\mathbf{v}_t = \mathbf{u}_t + \mathbf{B}\left(-\boldsymbol{\kappa}\mathbf{W}_t - \boldsymbol{\psi}\vartheta_t + \mathbf{A}'\mathbf{v}_t\right)$$

$$\mathbf{u}_t - \mathbf{B}\begin{bmatrix} \boldsymbol{\kappa} & \boldsymbol{\psi} & -\mathbf{A}' \end{bmatrix}\begin{bmatrix} \mathbf{w}_t \\ \vartheta_t \\ \mathbf{v}_t \end{bmatrix} = \mathbf{u}_t - \mathbf{B}\widehat{\mathbf{A}}'\mathcal{W}_t \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots .82$$

So inserting equation 82 in to equation 81 will produce

$$LL = -\frac{Tn}{2}\log(2\pi) - \frac{Tn}{2} - \frac{T}{2}\log\left|\frac{1}{T}\sum_{t=1}^{T}\left(\mathbf{u}_t + \mathbf{B}\widehat{\mathbf{A}}'\mathcal{W}_t\right)\left(\mathbf{u}_t + \mathbf{B}\widehat{\mathbf{A}}'\mathcal{W}_t\right)'\right| \dots\dots\dots .83$$

To find the coefficients of the above maximum likelihood equation lets define the two matrixes using $\mathbf{u}_t$ which is $n+2$ dimension vector and $\mathcal{W}_t$ which is also $n+2$ dimension vector

$$\left(\sum_{w_t w_t}\right)^{-1}\left(\sum_{w_t \mathbf{u}_t}\right)\left(\sum_{\mathbf{u}_t \mathbf{u}_t}\right)^{-1}\left(\sum_{\mathbf{u}_t w_t}\right) \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots .84$$

$$\left(\sum_{\mathbf{u}_t \mathbf{u}_t}\right)^{-1}\left(\sum_{\mathbf{u}_t w_t}\right)\left(\sum_{w_t w_t}\right)^{-1}\left(\sum_{w_t \mathbf{u}_t}\right) \dots\dots\dots\dots\dots\dots\dots .85$$

The Eigen value of the matrix in equation 84 when ordered from larger to smaller will enable us to identify the basis of the h cointegration relations. The bases of the h cointegration equation are equal to the Eigen vectors associated with the h largest Eigen values. But to identify the n Eigen vectors in matrix 84 we have to normalize them by $\mathbf{M}'\sum_{w_t w_t}\mathbf{M} = \mathbf{I}$ for $\mathbf{M}' = \begin{bmatrix} \mathbf{m}_1', \mathbf{m}_2', \mathbf{m}_3', \dots .\mathbf{m}_n' \end{bmatrix}$ when the vectors or $\mathbf{m}_i$ ordered in descending order of their Eigen values.

In other word the vector space of the cointegration equation $\left(\mathbf{A}\right)$ is spanned by the first h cointegration equations in $\mathbf{M}'$ or $\mathbf{A}' = \begin{bmatrix} \mathbf{m}_1', \mathbf{m}_2', \mathbf{m}_3', \dots .\mathbf{m}_h' \end{bmatrix}$. For matrix given in equation 85 similar Eigen vectors of $\mathbf{H}'$ can be generated and ordered in descending order in terms of their Eigen values, means $\mathbf{H}' = \begin{bmatrix} \mathbf{h}_1', \mathbf{h}_2', \mathbf{h}_3', \dots .\mathbf{h}_n' \end{bmatrix}$. But to uniquely identify the characteristic vectors they must be normalized by imposing $\mathbf{H}'\sum_{\mathbf{u}_t \mathbf{u}_t}\mathbf{H} = \mathbf{I}$.

As before if we generate $\mathbf{H}'\mathbf{u}_t = \boldsymbol{\rho}_t$ and $\mathbf{M}'\mathcal{W}_t = \boldsymbol{\tau}_t$ then it follows from the concept of canonical correlation that

$$E\left(\boldsymbol{\rho}_t\boldsymbol{\rho}_t\right) = \mathbf{I}, \ E\left(\boldsymbol{\tau}_t\boldsymbol{\tau}_t\right) = \mathbf{I} \text{ and } E\left(\boldsymbol{\rho}_t\boldsymbol{\tau}_t\right) = \mathbf{R} \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots .86$$

Following the same logic as before, the fully concentrated full information Johansen maximum log likelihood function is given as

$$LL = -\frac{Tn}{2}\log(2\pi) - \frac{Tn}{2} - \frac{T}{2}\log\left|\sum_{\mathbf{u}_t \mathbf{u}_t}\right| - \frac{T}{2}\log\prod_{i=1}^{h}\left(1 - \lambda_i\right) \dots\dots\dots\dots .87$$

And the coefficients in the vector error correction specified in equation 36 or 75 are derived in the following way

$$\zeta_i = \xi_i - \boldsymbol{\alpha}_h \varpi_i - \boldsymbol{\beta}_h \ell_i - \zeta_0 \chi_i \quad \text{for} \quad i = 0,1,2 \ldots P-1 \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 88$$

$$\boldsymbol{\alpha}_0 = \boldsymbol{\theta} - \boldsymbol{\alpha}_h - \boldsymbol{\beta}_h \gamma - \zeta_0 \wp \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 89$$

$$\mathbf{A}' = \left[\mathbf{m}_1', \mathbf{m}_2', \mathbf{m}_3', \ldots\ldots\mathbf{m}_h' \right] \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 90$$

$$\mathbf{B} = -\sum \mathbf{u}_t w_t \quad \widehat{\mathbf{A}} = \sum \mathbf{u}_t w_t \begin{bmatrix} \boldsymbol{\kappa} \\ \boldsymbol{\psi} \\ -\mathbf{A}' \end{bmatrix} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 91$$

$$\mathbf{B}\widehat{\mathbf{A}}' = -\sum_{\mathbf{u_t v_t}} \widehat{\mathbf{A}}\widehat{\mathbf{A}}' = \mathbf{B}\left[-\boldsymbol{\kappa} \quad -\boldsymbol{\psi} \quad \mathbf{A}' \right] = \left[-\mathbf{B}\boldsymbol{\kappa} \quad -\mathbf{B}\boldsymbol{\psi} \quad \mathbf{B}\mathbf{A}' \right]$$
$$= \left[-\mathbf{B}\boldsymbol{\kappa} \quad -\mathbf{B}\boldsymbol{\psi} \quad \mathbf{B}\mathbf{A}' \right] = \left[-\boldsymbol{\alpha}_h \quad -\boldsymbol{\beta}_h \quad -\zeta_0 \right] \qquad \ldots\ldots\ldots\ldots\ldots\ldots 92$$

$$\boldsymbol{\Omega} = \frac{1}{T} \sum_{t=1}^{T} \left(\mathbf{u}_t - \boldsymbol{\alpha}_h \mathbf{W}_t - \boldsymbol{\beta}_h \vartheta_t - \zeta_0 \mathbf{v}_t \right) \left(\mathbf{u}_t - \boldsymbol{\alpha}_h \mathbf{W}_t - \boldsymbol{\beta}_h \vartheta_t - \zeta_0 \mathbf{v}_t \right)' \ldots\ldots\ldots\ldots\ldots 93$$

These are the 5 versions of the Johansen full information maximum likelihood estimation models that can be used. The selection of a specific model can be based on some chi-square or $\chi^2(k)$ testes when ever possible. The chi-square test can be used by starting from the unrestricted to ward the more restricted versions (Johansen 1998 and 1991) given all are having the same lag order. And this will be very help full procedure if the conventional approach to cointegration vectors were followed. This is so because under the null that there are h cointegration equations the testes will follow conventional distributions or specifically standard normal and chi-square distributions (Johansen 1998 and 1991). So when ever it is possible to nest two models one over the other this test will be used to find out which model is more parsimonious. However in the search it is not possible to assume one model fits all since the price pattern and dynamics of log of transaction cost can be different from market to market.

So a pragmatic approach is to depend on visual inspection of the level data and the log of price margin between two market prices to select appropriate model. But as was stated before the most appropriate models based on theory are model 2, 4 and 5. The assumption of zero log of price margins is not logical for grains traded in spatially dispersed markets. Moreover the data does not have any clear deterministic linear trend let alone deterministic quadratic trend so the use of model one is not advisable. The most logical models are model 4 and 5 but in order to allow for seasonal dynamics with out allowing any form of trend in error correction or cointegration equations the inclusion of model 2 is also productive. In short model 2, 4 and 5 will be used in this paper to estimate long run relationships between prices of white wheat discovered in spatially dispersed markets.

2.5. <u>Estimation of the persistence statistics or profile</u>

In bivariate vector error correction model the size and sign of adjustment parameters and the statistical significance of the adjustment parameters will be used to analyze the short

run dynamics of the market. If adjustment parameter is 0.33, it means it will take 3 (= 1/0.33) periods for the market to correct the shock initiated in its long run relationship with other markets. Moreover if the markets have shorter lags it means the markets have short memory of past shocks or to follow Ravallion (1988) thinking they have strong cointegration in short run. To summarize the impact of both short run groups of parameters impulse response functions are widely used in literature. However impulse response are found to be ineffective, if shocks are correlated and the orthogonaization method used to solve the problem is found to be inconclusive, since it is order sensitive (Pesaran and Shin 1996).

Following Pesaran and Shin (1996) persistence profiles are used which are order insensitive to measure the persistence of system wide shock injected in to the cointegration vector equal to $\boldsymbol{\beta}'\boldsymbol{\Omega}\boldsymbol{\beta}$. Where $\boldsymbol{\beta}$ is the cointegrating vector and $\boldsymbol{\Omega}$ is the variance covariance matrix. The advantage of persistence profile is that it is order insensitive for given identification assumption imposed in error correction model. However still it is sensitive to the identification assumption imposed. The identification assumption in this paper is selected based on level of normality of the error vectors.

Following shock equal to $\boldsymbol{\beta}'\boldsymbol{\Omega}\boldsymbol{\beta}$ at period zero the change in variance of forecast error in period n is given by $\boldsymbol{\beta}'\boldsymbol{\beta}_n\boldsymbol{\Omega}\boldsymbol{\beta}_n'\boldsymbol{\beta}$ where $\boldsymbol{\beta}_n$ is equal to

$$\boldsymbol{\beta}_n = \boldsymbol{\Phi}_1\boldsymbol{\beta}_{n-1} + \boldsymbol{\Phi}_2\boldsymbol{\beta}_{n-2} + \boldsymbol{\Phi}_3\boldsymbol{\beta}_{n-3} + ...\boldsymbol{\Phi}_p\boldsymbol{\beta}_{n-p}$$ …………………………………………………94

Given the matrix $\boldsymbol{\Phi}_i$ is as defined in equation 26. At limit given $\boldsymbol{\beta}$ is cointegrating vector $\boldsymbol{\beta}'\boldsymbol{\beta}_n\boldsymbol{\Omega}\boldsymbol{\beta}_n'\boldsymbol{\beta}$ will approach zero. So in persistence profile the adjustment process following shock equal to $\boldsymbol{\beta}'\boldsymbol{\Omega}\boldsymbol{\beta}$ can be analyzed by considering the dynamics of relative persistence given by

$$\mathbf{h}(n) = \boldsymbol{\beta}'\boldsymbol{\beta}_n\boldsymbol{\Omega}\boldsymbol{\beta}_n'\boldsymbol{\beta} \times \mathbf{Diag}\left(\left(\boldsymbol{\beta}'\boldsymbol{\Omega}\boldsymbol{\beta}\right)^{-1}\right)$$ …………………………………………………….95

The value of $\mathbf{h}(n)$ will range from 1 when $n = 0$ to 0 when $n = \infty$. The value at given period of n or $\mathbf{h}(n)$ will measure fraction of the initial shock which is persisting at that period. For example if $\mathbf{h}(2) = 0.3$, it means 30% of the shock is not corrected at second period or similarly 80% of the disequilibrium is corrected at second period.

2.6. <u>Estimation of the single common integrating trend</u>

If n markets are under rule of one price there will be $n - 1 = h$ cointegration relations and 1 common trend. The estimation of this single common trend will be useful in order to understand the importance of a given market in determination of the common trend. A market which is highly cointegrated with other markets, which is having significant impact on common trend and which is more or less weakly exogenous will be the market where efficient stabilization can be done in cost effective manure.

The estimation of the long run trend will be done using Gonzalo and Granger (1995) linear decomposition of price vector in permanent and temporary component as

$$\mathbf{P}_t = \mathbf{A}_1 f_t + \tilde{\mathbf{P}}_t \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots96$$

Where $\mathbf{P}_t$, $\mathbf{A}_1$ and $\tilde{\mathbf{P}}_t$ are n dimension vectors and f_t is a scalar of common trend related to single common trend under the model used here. So the prices are function of permanent component (f_t) loaded by loading matrix $(\mathbf{A}_1)$ plus temporary component $(\tilde{\mathbf{P}}_t)$.

The basic assumption imposed by Gonzalo and Granger (1995) and used to estimate the common trend is that f is linear on observed prices and the temporary component does not have permanent impact on prices. Formally the first assumption imply

$$f_t = \mathbf{a}_\perp^{'} \mathbf{P}_t \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots97$$

And the only definition of f_t which can grant the second assumption is found in null vector of the cointegration vectors related to Eigen vectors of matrix in equation 43, 68 or 85 or simply the null vector of $\mathbf{A}$ (cointegrating matrix) in $\mathbf{H}^{'} = \left[\mathbf{h}_1^{'}, \mathbf{h}_2^{'}, \mathbf{h}_3^{'}......\mathbf{h}_{n1}^{'}\right]$ given by

$$\mathbf{H}^{'} = \left[\mathbf{h}_1^{'}, \mathbf{h}_2^{'}, \mathbf{h}_3^{'}......\mathbf{h}_n^{'}\right] = \left[\mathbf{h}_1^{'}, \mathbf{h}_2^{'}, \mathbf{h}_3^{'}......\mathbf{h}_h^{'}, \mathbf{h}_{n=h+1}^{'}\right] = \left[\mathbf{h}_1^{'}, \mathbf{h}_2^{'}, \mathbf{h}_3^{'}......\mathbf{h}_h^{'}, \mathbf{a}_\perp^{'}\right] \dots\dots\dots\dots98$$

The first vectors related to the h largest Eigen values are the vectors which are the dual of the cointegrating vectors. So they can't be orthogonal with them. The remaining one vector will span the null space of cointegrating vectors or it is orthogonal to the cointegrating space. Formally

$$\mathbf{h}_n^{'} \mathbf{m}_i = \mathbf{a}_\perp^{'} \mathbf{m}_i = 0 \qquad \text{for } i = 1, 2, 3, ... h \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots99$$

So the common trend will be given by equation 97. The statistical significance and the numeric significance of $\mathbf{a}_\perp^{'}$ will be use full in providing information about the relative importance of a given market in the general price formation. And as proved by Gonzalo and Granger (1995) a conventional wald test can be applied given the vector error correction model is correctly specified, identified and estimated by Johansen (1988 and 1991) frame work. Or simply we can use the following log likelihood test

$$L = -T \sum_{i=r+1}^{P} \ln\left((1-\lambda_{P-i})/(1-\lambda_P)\right) \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots100$$

Where λ_p is the smallest Eigen value in the unrestricted model and λ_{p-i} is the smallest Eigen value when the impact of i markets is constrained to be zero. This is has Chi2 distribution with $p-m$ degree of freedom. Where p and m is number of Eigen values in unrestricted and restricted versions, respectively. One important point is that if eigen

vector is identified by $\widehat{\mathbf{H}}^{'}\widehat{\mathbf{H}} = \mathbf{I}$ than $\mathbf{H}^{'}\sum_{w_t w_t}\mathbf{H} = \mathbf{I}$, it can be changed from $\widehat{\mathbf{H}}$ to $\mathbf{H}$ by using the formula $\mathbf{h}_i = \widehat{\mathbf{h}}_i \div \sqrt{\mathbf{h}_i^{'}\sum_{w_t w_t}\mathbf{h}_i}$ (Hamilton, 1994)

2.7. <u>Testes used to verify the specification of the error correction model</u>

In this paper test for normality is done by using Jarque and Bera (1980 and 1981) test for normality, Lagrange multiplier test developed by Breusch (1978) and Godfrey (1978) is used for serial correlation and trace statistics and maximum Eigen value developed by Johansen (1988 and 1991) are used to find the rank of the cointegration vectors. ARCH effect is tasted following Engle (1982) recommendation.

Since the Johansen full information maximum likelihood estimation assumes that the error vector follows a multivariate normal distribution with out any serial correlation and constant variance (Johansen 1998 and 1991 and Hamilton 1994) these tests are critical for correct specification of the model. Moreover the unit root testes used in this paper or the ADF test is based on standard Brownian distribution which is adapted to the test under the assumption of white noise error terms (Dickey and Fuller 1979 and Hamilton 1994). So the white noise-ness testes are as critical for unit root testes as they are for the vector error correction model. Let's analyze each of the testes one after the other below

2.7.1. <u>Normality test</u>

To fully grasp the Jarque and Bera (1980 and 1981) or simply JB test let's follow theoretical exposition given by Davidson and Mackinnon (1993) below. For error vector $\mathbf{\varepsilon}_t$ in one of the error correction models stated from equation 32 to equation 36, ε_{ti} as one scalar in $\mathbf{\varepsilon}_t$ will be related to price of market i. Under the null hypothesis of that ε_{ti} is normally distributed white noise process, the following asymptotic limits are true.

$$\lim_{T\to\infty}\left(\frac{1}{T}\sum_{t=1}^{T}\left(\frac{\varepsilon_{ti}}{\sigma^2}\right)^2\right) = 6 \dots\dots\dots\dots101$$

$$\lim_{T\to\infty}\left(\frac{1}{T}\sum_{t=1}^{T}\left(\frac{\varepsilon_{ti}}{\sigma^4}-3\right)^2\right) = 24 \dots\dots\dots\dots102$$

For[8] $E(\varepsilon_{ti}) = \hat{\varepsilon}_{ti}$ the following standard normal value (Z_{ti}) can be defined as $Z_{ti} = (\varepsilon_{ti} - \hat{\varepsilon}_{ti})/\sigma$. Under the null of normality, the following statistics for Skewness has chi-square distribution with 1 degree of freedom or it follows $\chi^2(1)$ distribution with mean 1 and variance of 2.

---

[8] which can be different from zero for vector error correction model with out unrestricted drift in the main model specified in model 3 and 4

$$\frac{1}{6T}\left(\sum_{t=1}^{T}Z_{ti}^{3}\right)^{2} = S \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots 103$$

Under assumption of normality the associated statistics for kurtosis is given by

$$\frac{1}{24T}\left(\sum_{t=1}^{T}Z_{ti}^{4}-3\right)^{2} = K \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots 104$$

This also has chi-square distribution with 1 degree of freedom or it follows $\chi^2(1)$ distribution with mean 1 and variance of 2. The Jarque and Bera (JB) test for normality is developed by combining the Skewness and kurtosis test in following manner

$$\frac{1}{6T}\left(\sum_{t=1}^{T}Z_{ti}^{3}\right)^{2} + \frac{1}{24T}\left(\sum_{t=1}^{T}Z_{ti}^{4}-3\right)^{2} = N \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots 105$$

This has chi-square distribution with 2 degree of freedom or follows $\chi^2(2)$ distribution with mean of 2 and variance of 4. And for n error terms in error correction model the vector level normality can be tested by

$$\sum_{j=1}^{n}S_{j} \ , \ \sum_{j=1}^{n}K_{j} \ \text{and} \ \sum_{j=1}^{n}N_{j} \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots 106$$

For testing the null of zero skewness, zero kurtosis and over all normality, respectively. The skewness and kurtosis testes will follow $\chi^2(n)$ or χ^2 distribution with n degree of freedom and the normality test will follow $\chi^2(2n)$ or χ^2 distribution with $2n$ degree of freedom.

One important point is that in large samples the assumption of normality is not binding for vector error correction model. The cointegration analysis is asymptotically applicable for both Gaussian and non Gaussian distribution (Johansen 1991 and Hamilton 1994) but given moderately small sample size of 96 months used in this paper depending on limit distribution to make inference is not optional. So normality is demanded in model selection; by making sure that the null of normality is not rejected up to 10% level of significance, than the conventional 5% level.

For unit root test the limit distributions are not sensitive to normality but in small samples normality is critical. So the same test specified in equation 104, 105 and 106 are used in the residuals of the regression specified in equation 7. And still 10% significance than 5% significance is used in model selection.

Even though Jarque and Bera (JB) test for normality is extensively used in stationary series or VAR, a Monte Carlo simulation by Demiroglu (2000) did show that it is equally applicable to both integrated series and cointegrated vectors. So the use of Jarque and Bera (JB) test to test normality for error vectors of vector error correction model is logically sound.

2.7.2. <u>Test for serial correlation</u>

To test for serial correlation Breusch – Godfrey (BG) or Lagrange-multiplier (LM) test developed by Breusch (1978) and Godfrey (1978) is used in all cases. For unit root test the assumption of no serial correlation is fundamental for using the standard Brownian distribution which is the building block of the ADF tests. So the existence of serial correlation is tested by using the conventional Lagrange-multiplier test from 0 up to 15 lags to account for any form of seasonality. The selection of the specific lag will be dependent on both information criterions and the white nosiness of the error terms. To explain the conventional approach let's reproduce equation 7 again

$$\Delta P_t = \alpha_t + \beta t + \varsigma_1 \Delta P_{t-1} + \varsigma_2 \Delta P_{t-2} + ... + \varsigma_{m-1}\Delta P_{t-m+1} + \lambda P_{t-1} + \varepsilon_t \ ....7$$

So it is assumed that the error term in equation 7 (ε_t) will follow auto regressive process up to m lag. Formally $\varepsilon_t = AR(m)$ or

$$\varepsilon_t = \sigma_1 \varepsilon_{t-1} + \sigma_2 \varepsilon_{t-2} + \sigma_3 \varepsilon_{t-3} + ... \ ... + \sigma_m \varepsilon_{t-m} + \mathbf{bX} + e_t \ ....107$$

e_t is assumed to follow a white noise process. The vector $\mathbf{X}$ is related to all explanatory variables in equation 7. And for T total observations used with P lags in the main regression only $(T-P)$ error terms are available. So for R^2 which is degree of determination in equation 107 $(T-P-m)R^2$ will follow $\chi^2(m)$ with mean m and variance of $2m$ under the null $\sigma_i = 0$ for all i from one to m. For the vector error correction model let's use the unrestricted error correction model given in equation 32 to nest the five models as replicated below

$$\Delta \mathbf{P}_t = \mathbf{\alpha} + \mathbf{\beta}t + \zeta_1 \Delta \mathbf{P}_{t-1} + \zeta_2 \Delta \mathbf{P}_{t-2} + ... + \zeta_{P-1}\Delta \mathbf{P}_{t-p+1} + \zeta_0 \mathbf{P}_{t-1} + \mathbf{\varepsilon}_t \ ....32$$

Given $\mathbf{P}_t$ is n dimensional vector. If under the alternative hypothesis of m order vector serial correlation of the error term we generate the following auxiliary VAR on error vector $(\mathbf{\varepsilon}_t)$

$$\mathbf{\varepsilon}_t = \mathbf{\Omega}_1 \mathbf{\varepsilon}_{t-1} + \mathbf{\Omega}_2 \mathbf{\varepsilon}_{t-2} + \mathbf{\Omega}_3 \mathbf{\varepsilon}_{t-3} + ... + \mathbf{\Omega}_m \mathbf{\varepsilon}_{t-m} + \mathbf{BX} + \mathbf{e} \ ....108$$

Again $\mathbf{X}$ as defined above and the unrestricted version of equation 32 can be written as

$$\Delta \mathbf{P}_t = \mathbf{\alpha} + \mathbf{\beta}t + \sum_{i=1}^{p-1} \zeta_i \Delta \mathbf{P}_{t-i} + \sum_{i=1}^{m} \mathbf{\Omega}_i \mathbf{\varepsilon}_{t-i} + \mathbf{e} \ ....109$$

In this paper P is defined as lags of the error correction model or simply lags and m is defined as serial lag to avoid confusion. And $\mathbf{e}$ is a white noise error vector or it is identically and independently distributed multivariate normal vector. And as shown by Brüggermann et al (2006) the different specification implemented in the error correction model will not make difference on the critical value of the limit distribution. So we can use the following LM statistics which is similar to the conventional LM statistics applied

to VAR model.

$$\frac{1}{T}\mathbf{C_k}'\left[\sum\nolimits_{\mathbf{uu}}\right]^{-1}\mathbf{C_k}$$ 110

For $\sum_{\mathbf{uu}}$ as defined in equation 42 and 43 or 67 and 78 or 84 and 85, and $\mathbf{C}_k$ is the matrix created by stacking k vectors of auto correlation coefficients $(\mathbf{c}_i)$ of the error term. The vector $\mathbf{c}_i$ is equal to the i lag autocorrelation of log price vector. The statistics in equation 110 will follow $\chi^2(mn^2)$ distribution with mean of mn^2 and variance of $2mn^2$ (ibid). Note that m is equal to the lag used in serial correlation test and n is the number of prices under consideration.

The soft ware used in this paper, Stata 9, fits a different but related statistics. For c_i which is i^{th} order auto correlation coefficient or the ith scalar in $\mathbf{C_k}$ the statistics used to test for serial correlation by Stata 9 is given below

$$\frac{1}{T}c_i'\left[\sum\nolimits_{\mathbf{uu}}\right]^{-1}c_i$$ 111

The statistics in equation 111 will follow $\chi^2(n^2)$ distribution with mean of n^2 and variance of $2n^2$ Brüggermann et al (2006). So the existence of serial correlation up to 15th lag will be tested by considering 1 to 15 lags, separately. In this paper to avoid ambiguity the statistics in equation 110 will be defined as statistics for over all serial correlation or cumulative serial correlation and the statistics in equation 111 will be defined as simple serial correlation of a given lag. As proofed by Brüggermann et al (2006) the imposition of any of the VECM restrictions or inclusion of centered seasonal dummies will not effect the limit distribution and the test result.

The standard Brownian motion used to derive the cointegration rank is dependent on assumption of independently distributed error terms and the assumption is as critical for limit distributions as for small samples (Johansen 1991 and Hamilton 1994). So the test for serial correlation is taken seriously in each stage of the search for $n-1$ cointegrating prices under one common trend. And again strict 10% level of significance is adhered than the conventional 5% level of significance.

2.7.3. Test for vector rank or number of cointegration equations

For testing the rank of the vector error correction model or the number of cointegration vectors that have to be fitted in the error correction model, Johansen (1988 and 1991) trace statistics and Maximum Eigen value are used. These tests are simple extension of the full information maximum log likelihood identified for model 1 to 3 in equation 58, for model 4 in equation 70 and for model 5 in equation 87. Now let's replicate them here under null hypothesis that there are just h cointegration equations in VEC model or LL_H

$$LL_H = -\frac{Tn}{2}\log(2\pi) - \frac{Tn}{2} - \frac{T}{2}\log\left|\Sigma_{\mathbf{u_t u_t}}\right| - \frac{T}{2}\log\prod_{i=1}^{h}(1-\lambda_i) \quad \text{................112}$$

The alternative hypothesis is that there are $h+1$ cointegration equations. So the log likelihood under alternative hypothesis is

$$LL_A = -\frac{Tn}{2}\log(2\pi) - \frac{Tn}{2} - \frac{T}{2}\log\left|\Sigma_{\mathbf{u_t u_t}}\right| - \frac{T}{2}\log\prod_{i=1}^{h+1}(1-\lambda_i) \quad \text{.................113}$$

So two times the difference between the log likelihood under the alternative and the null will follow none standard distribution developed by using Brownian distribution and critical values are tabulated by Johansen (1988 and 1991). This is called maximum Eigen value or statistics, since it is calculated by considering the next larger Eigen value following the largest h Eigen values or by analyzing λ_{h+1}. Formally

$$\lambda_{\max} = 2(LL_A - LL_H) = -T\log(1-\lambda_{h+1}) \quad \text{..114}$$

As stated in Verbeek (2004) the idea behind the test can be grasped by considering equation 112 in which the largest h Eigen values are used in the model, since they are related to the h largest canonical correlations. And other Eigen values are fixed at 0. So the above test will check, if the next largest $h+1$ Eigen value is different from zero. Means it will test, if there is one additional cointegrating vector in the model or not.

Another alternative test is to use trace statistics in which the null is given under equation 112 with h cointegration equations against the alternative hypothesis which assumes that there are n cointegration equations. Means the sample space spanned by the cointegration vectors is n. This is like claming all of them are stationary. The log likelihood under alternative hypothesis will be

$$LL_A = -\frac{Tn}{2}\log(2\pi) - \frac{Tn}{2} - \frac{T}{2}\log\left|\Sigma_{\mathbf{u_t u_t}}\right| - \frac{T}{2}\log\prod_{i=1}^{n}(1-\lambda_i) \quad \text{....................115}$$

The associated trace statistics which checks if the Eigen values from $h+1$ to n collectively are equal to zero or not is given by the following trace statistics.

$$\lambda_{\text{trace}} = 2(LL_A - LL_H) = -T\log\prod_{i=h+1}^{n}(1-\lambda_i) \quad \text{.......................................116}$$

The standard Brownian motion used to generate the limit of these two statistics depends on assumptions made about the existence or non existence of drift and trend terms in the data generating process. Means for the 5 models given from equation 32 to 36 there are 5 different critical values tabulated for each model specification.

The question is which one of the above testes to use? Monte Carlo studies show that both are imperfect testes and are not superior to one another. In this paper trace test is used for the search process but for those markets under rule of one price based on trace test a maximum Eigen value based test is also supplemented to see if there is inconsistency in result of both testes.

2.7.4. <u>Test for time varying heteroskedasticity or ARCH effect</u>

The Johansen vector error correction model and related rank test are asymptotically applicable for both homoskedastic and heteroskedastic errors (Johansen 1988 and 1991, Hamilton 1994, Cavaliere et al 2009 and Lee and Tse 1996). However in small and moderately small samples heteroskedasticity is observed to reduce the power of rank testes marginally (Lee and Tse 1996 and Cavaliere et al 2009).

To test for auto regressive conditional heteroskedasticity (ARCH) effect discovered by Engle (1982) or its extension to generalized auto regressive conditional heteroskedasticity (GARCH) by Bollerslev (1986) we can use the residuals from the vector error correction model. Let's start from the following VECM given in equation 32 and replicated below

$$\Delta \mathbf{P}_t = \boldsymbol{\alpha} + \boldsymbol{\beta} t + \boldsymbol{\zeta}_1 \Delta \mathbf{P}_{t-1} + \boldsymbol{\zeta}_2 \Delta \mathbf{P}_{t-2} + ... + \boldsymbol{\zeta}_{P-1} \Delta \mathbf{P}_{t-p+1} + \boldsymbol{\zeta}_0 \mathbf{P}_{t-1} + \boldsymbol{\varepsilon}_t \dots\dots\dots\dots\dots\dots\dots..32$$

When this is generalized to VECM with GARCH effect it will be

$$\Delta \mathbf{P}_t = \boldsymbol{\alpha} + \boldsymbol{\beta} t + \boldsymbol{\zeta}_0 \mathbf{P}_{t-1} + \sum_{i=1}^{p-1} \boldsymbol{\zeta}_i \Delta \mathbf{P}_{t-i} + \mathbf{e}\left(\sqrt{ \boldsymbol{\sigma}^2 + \sum_{i=1}^{m} \boldsymbol{\Psi}_i \boldsymbol{\sigma}^2_{t-i} + \sum_{i=1}^{n} \boldsymbol{\Omega}_i \boldsymbol{\varepsilon}^2_{t-i} } \right)$$

$$\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots..117$$

In the above equation **e** is standard normal white noise vector. And equation 117 will be equal to equation 32 if $\sum_{i=1}^{m} \boldsymbol{\Psi}_i = \sum_{i=1}^{n} \boldsymbol{\Omega}_i = \mathbf{0}$ is true. The test for GARCH (m, n) is equal to test of ARCH(m + n) and following Engle (1982) and Bollerslev (1986) the following auxiliary regression can be used to test if $\sum_{i=1}^{m} \boldsymbol{\Psi}_i = \sum_{i=1}^{n} \boldsymbol{\Omega}_i = \mathbf{0}$ is holding for each error term i assuming that $\boldsymbol{\Psi}_i$ and $\boldsymbol{\Omega}_i$ are block diagonal[9].

$$\varepsilon^2_{it} = \phi_0 + \sum_{j=1}^{m+n} \phi_j \varepsilon^2_{it-j} + e \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots...118$$

Given $\phi_0 > 0$ and $\phi_j > 0$ for all j and $\phi_0 + \sum_{j=1}^{m+n} \phi_j < 1$ the test will follow conventional normal, F, student T and Chi-square destructions under the null the errors in restricted version of equation 32 are white noise. For convince simple OLS result will be used to test for ARCH effect since it is best unbiased estimator among linear estimators (BLUE). How ever if the restrictions are valuated in OLS restricted maximum likelihood regression will be used, which is best among all possible estimators (BUE).

---

[9] If not there is need for using multivariate ARCH or GRACH test which can complicate the analysis.

Part three
3. Discussion and analysis
3.1. Unit root test for 8 white wheat whole sale prices

The building blocks of cointegration analysis are variables which are generated by none stationary data generating process. So the first step in the cointegration analysis is to separate markets which are following unit root from those markets which are stationary. In this subpart of the paper the data generating process of the sample data used is tested by using ADF version of t and F testes to see if it have unit root. However there are two problems in practical application of the method. First there is need to identify the right lag length to augment the standard DF test. Second the right test has to be used to make sure that the test used is consistent with the data generating process which generated the sample the data.

Figure 3 levels of 8 white wheat whole sale log prices in Ethiopia from 1996 to 2003 G.C.
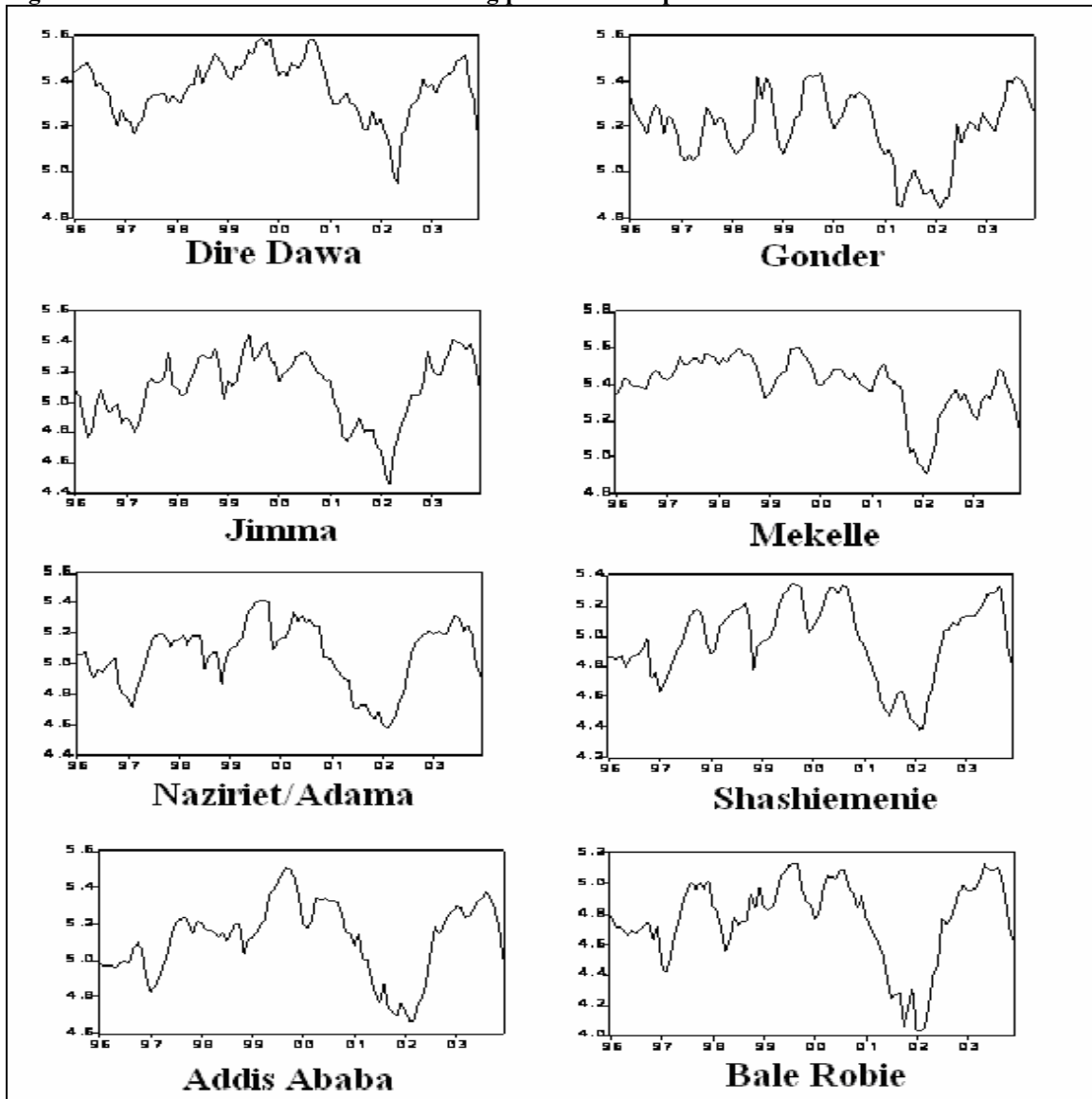
Lets start from the second problem of using the right version of ADF test from ADF with out drift (t_{df}), ADF with drift (t_{α}) or ADF with deterministic trend (t_T). There are two complementary ways to address the problem. One is to use graphic presentation of the level data. In which t_T will be appropriate if there is trend in level data. And as can be seen from figure one above there is no trend in level data and this test is less appropriate. The ADF test with drift or t_{α} test will be appropriate if there is none zero average on the level data which can be caused by unit root around zero or by stationary series around drift. And this is appropriate test to start the test process since the level data is having none zero average. However if the data actually have no drift on it ADF test with out drift or t_{df} is more power full than ADF test with drift or t_{α} .

The problem with these testes is that, they highly depended on subjective judgment of the researcher. The alternative is to use F type of testes developed by Dickey and Fuller (1979) and termed as Φ_1 and Φ_3 by Peterson (2000). The F version test of Φ_1 will jointly test if the drift term is zero and there is unit root in the data. And Φ_3 will jointly test if the trend term is zero and there is unit root in the data. But the problem with these two testes is related to the fact that they have weak power, especially if they are different from the real data generating process (ibid). But most importantly the alternative hypothesis is related to the fact that either the data is stationary or there is drift/trend or both. As result rejection of the null will not give us a conclusive result. The best way is to follow Peterson (2000) recommendation and mix both methods. Use the F version testes but their conclusion when the null is rejected must be evaluated in reference to the data generating process observed in the graphic presentation of the data. And following the result of the F version testes the more power full student t kind of testes has to be used to consolidate the conclusion.

The second problem is related to the fixation of appropriate lag. If lower lags are used the data may have serial correlation which can distort the conclusion. If higher lags are used ADF test will have very weak power. So the appropriate lag order must be selected first in order to reach sound conclusion. For this purposes four information criterions are used. These are Hannan and Quinn information criterion (HQIC), Final prediction error (FPE), Schwartz's Bayesian information criterion (SBIC) and Akaike's information criterion (AIC). The problem is that these information criterions are observed to pick different lags and there is no strong evidence to select one information criterions over the other[10]. AIC and FPE have higher probability of over fitting on lower dimensions but HQIC and mainly SBIC have higher probability of under fitting. And these problems are series in small or moderately small samples sizes. To complicate things if there is moving average component in the data, it is observed that all will under fit the right parsimonious auto regressive model. In this paper the result of these information criterions is considered as the best educated guess under some assumptions employed in penalty function of the information criterions than the true lag selection. And the lowest lag selected by any of the information criterions will be used as long as the error terms are independent. But if

---

[10] The Monte Carlo evidence will be presented latter on

serial correlation is observed lags will be increased until the error term becomes independent[11]. The upper limit for these lag augmentation is the maximum lag selected by any of the information criterions. This is done since independence of the error terms is very critical assumption for the critical values developed by Dickey and Fuller (1979). And given monthly nature of the data the maximum lag allowed in the information criterions is 15 months. Now lets discus the result.

Table 1 ADF test for level data of 8 wheat markets based lag selected by information criterions and independence of error terms

Market	With out drift		With drift			With trend		
	Lag used	t_{df}	Lag used	t_α	Φ_1	Lag used	t_T	Φ_3
Addis Ababa (A)	1	-0.121	1	-2.235	6.94***	1	-2.179	2.6
Bale Robe (B)	0	-0.282	0	-1.798	1.63	0	-1.798	1.65
Dire Dawa (D)	1	-0.607	1	-2.078	2.32	0	-1.92	1.97
Gonder (G)	5	-0.032	3	-2.799*	3.92	1	-2.71	3.72
Jimma (J)	0	0.04	0	-2.037	2.08	0	-2.036	2.07
Mekelle (M)	5	-0.395	2	-2.366	2.92	1	-3.221*	5.44
Nazareth (N)	4	-0.164	3	-3.312**	3.33	0	-1.848	1.75
Shashimiene (S)	1	0.203	1	2.51	3.15	1	-2.478	3.2

Note 1 * Significant at 10%, ** significant at 5% and * significant at 1% level.**

The joint test of the trend coefficient and the unit root coefficient under Φ_3 shows that the null of unit root around drift can't be rejected for all eight markets. However the conclusion can't be accepted at its face value since the test is not consistent with data generating process and serial correlation was observed in all possible lags. When there is serial correlation the DF distribution used to make inference is not appropriate and when the test is not consistent with the data generating process it have weak power to reject the null. One way or another, the conclusions have to be accepted with great reservation. However assuming there is trend in the data and t_T is appropriate test the null of unit root around trend is rejected at 10% but not at 5% level for Mekelle. If this test result was right there should be deterministic trend in level of Mekelle, which is not the case. So we have to reject this test result, too.

When the Φ_1 version of the F test is used the null of unit root around zero is not rejected even at 10% level for all but Addis Ababa. For Addis Ababa the null is rejected at 1% level which could imply that there is either unit root around drift, drift with stationery series or unit root around drift. If there is unit root around drift there must be stochastic trend in level data. However figure 3 above clearly shows that there is no trend in level of Addis Ababa. The second option is that the data is stationary around drift but the more power full test of t_α under such assumption is not rejecting unit root around drift at 10% level. So the only logical conclusion is that the data is following unit root around zero.

---

[11] Serial correlation up to 15th lag and JB normality test are performed but are not given here to reduce the number of tables

But under such assumption both Φ_1 and t_α have weak power so it have to be proven by more power full ADF test of t_{df} .

Assuming there is drift in the level data of Gonder at 10% and Nazareth at 5% level are having stationary series. But since the existence of drift is rejected by Φ_1 test these result can't be accepted. So as can be judged from the visual inspection of figure 3 and the more general F version testes the most appropriate test is ADF with out drift or t_{df} . And under this test the null of unit around zero can't be rejected at 10% level for all markets. There fore the most logical conclusion is that all prices are having at least one unit root.

The above conclusion is based on independently distributed error terms which are not necessarily normal. And ADF test is asymptotically applicable to all independent distributed errors (Dickey and Fuller 1979). However in small and moderately small samples used in this paper the normality of the error vectors is very critical for ADF testes (Hamilton 1994). For markets which are not having white noise error term in table 1 above a higher lags are selected based on white noisiness of the error terms, as given in table 2 below. Since when deterministic trend is allowed the error terms for all markets are not becoming white noise for any lag used in the ADF test with trend it is dropped from the analysis. Moreover for Shashimiene it was not possible to find white noise error terms given the fact that normality and independence are observed to take turn on different lags but never happen together. So Shashimiene is also dropped. One problem with using higher lags to create white noise error terms is related to the fact that the testes will have weak power to reject false null of unit root.

Table 2 ADF test for level data of 8 wheat markets lag selected by white nosiness of error terms

Market	With out drift		With drift		
	Lag used	t_{df}	Lag used	t_α	Φ_1
Addis Ababa (A)			0	-1.651	1.36
Dire Dawa (D)	26	-0.517	26	-0.318	0.18
Gonder (G)	32	-0.441	32	-2.093	2.29
Nazareth (N)	23	-0.382	19	-1.845	0.83

Note 2 * Significant at 10%, ** significant at 5% and * significant at 1% level.**

As can e seen from table 2 above all markets are having unit root around zero even when longer lags are used to find white noise errors. This is expectable since ADF testes have weak power when longer lags are used. Given there is no strong evidence to the contrary this analysis will continue by assuming that there is at least one unit root in all prices. However there is need to test the existence second unit root in the data.

Even though two unit roots are less common in economics the existence of two unit roots can be tested by using the first difference as level data and making the same analysis on first difference of the data. But for better conclusion, let's observe the graphic presentation of the pattern of first difference of level price given in figure 2, below.

Figure 4 First difference of 8 white wheat whole sale prices in Ethiopia from 1996 to 2003 G.C.
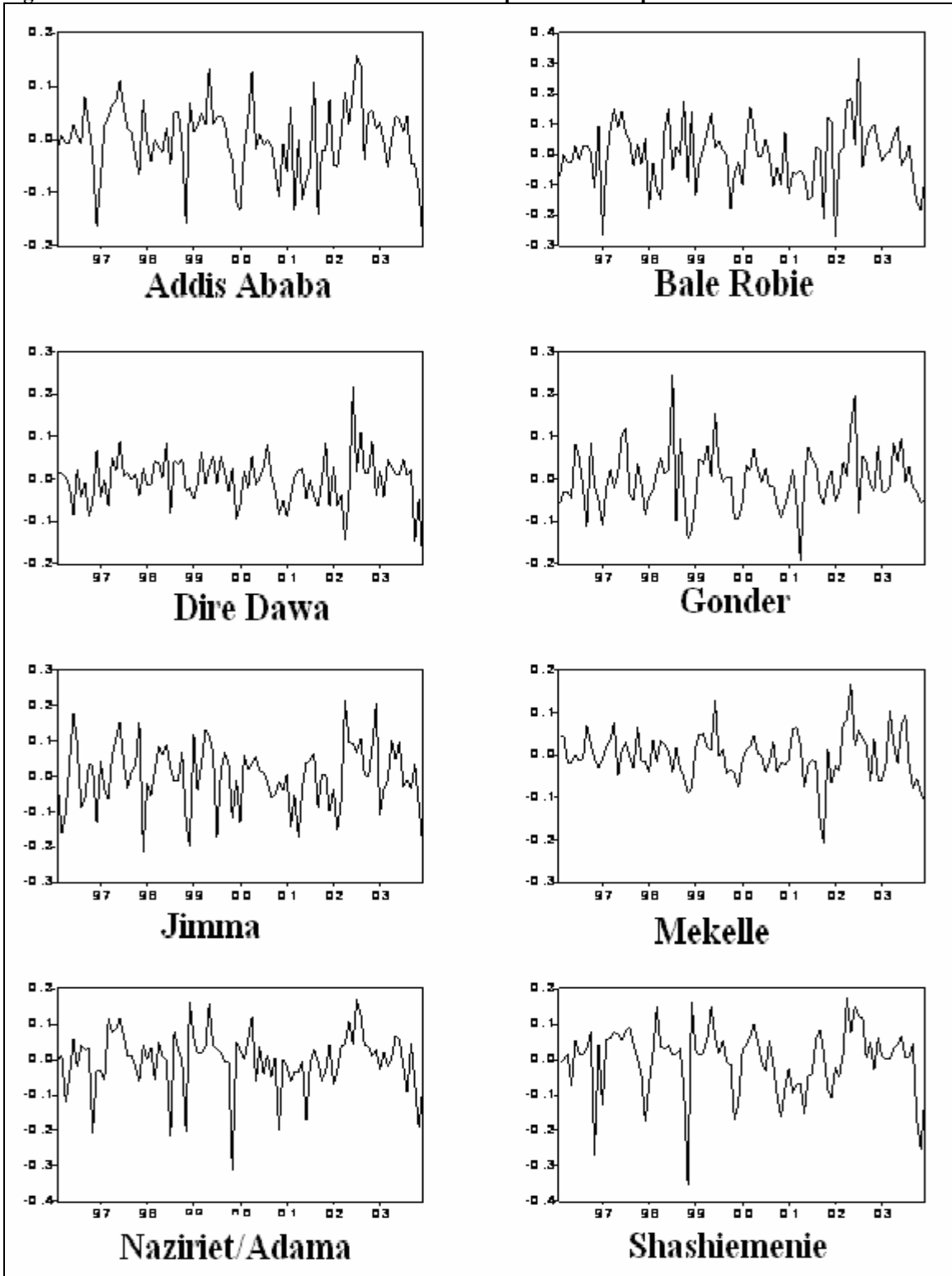
And as can be seen from figure 2 above the data is more of stationary around zero. But most importantly there is neither deterministic nor stochastic trend on it. This means the

appropriate test is ADF with out drift or t_{df}. But let's see if the testes can back the above visual conclusion.

Table 3 ADF test for first difference of level data of 8 wheat markets based on lag selected by information criterions and independence of error terms

Market	With out drift		With drift		
	Lag used	t_{df}	Lag used	t_{α}	Φ_1
Addis Ababa (A)	3	-3.409***	3	-3.369**	5.75**
Bale Robe (B)	3	-3.604***	2	-4.08***	8.35***
Dire Dawa (D)	1	-4.739***	1	-4.736***	11.31***
Gonder (G)	4	-5.192***	4	-5.157***	13.32***
Jimma (J)	3	-4.806***	3	-4.79***	11.48***
Mekelle (M)	4	-4.745***	4	-4.721***	11.2***
Nazareth (N)	3	-3.334***	3	-3.312**	5.5**
Shashimiene (S)	2	-4.864***	1	-4.833***	11.71***

Note 3 * Significant at 10%, ** significant at 5% and * significant at 1% level.**

Since the first difference does not have any trend Φ_3 is not appropriate but most importantly for all markets when trend is allowed it was not possible to find independent error terms at what ever lag. So it is dropped from the analysis. As can be seen from table 3 above the null of unit root around zero is rejected at 1% level for all markets except two. For these 6 markets the null of unit root with drift or with out drift is also rejected with 99% confidence. Means for these 6 markets there is no second unit root. For Addis Ababa and Nazareth the null of unit root around zero under Φ_1 and unit root around drift under t_{α} is rejected at 5% level but not at 10% level. And the null of unit root around zero based t_{df} is rejected at 1% level. However given the data is having zero mean it is known fact that t_{α} and Φ_1 have weak power compared to t_{df}, so these markets are also accepted to have only one unit root, too.

But again for Dire Dawa, Gonder and Nazareth even though there was no serial correlation in the above lags there is normality problem. So lags are added to find white noise error terms but it is important to remember that when the test is having longer lags it will have weak power to reject false null hypothesis.

Table 4 Unit root test for lag selected by distributional assumption for first difference

Market	With out drift		With drift		
	Lag used	t_{df}	Lag used	t_{α}	Φ_1
Dire Dawa	25	-1.448	25	-1.496	1.16
Gonder	31	-1.295	31	-1.345	0.91
Nazareth	22	-2.657***	22	-2.639*	3.53

Note 4 * Significant at 10%, ** significant at 5% and * significant at 1% level.**

And as expected when higher lags are used the test is having very weak power. The null of unit root around zero is not rejected for the first difference of all 3 prices up to 10% level. And the null of unit root around drift is not rejected at 10% level for 2 markets and at 5% level for Nazareth. And observing the pattern of the data at figure 2 above it is illogical to accept this conclusion which implies there is trend in the first difference data. When the more appropriate model of ADF with out drift is used Nazareth is having stationery fist difference of price at 1% level. However for Dire Dawa and Gonder which are using 25 and higher lags the null of unit root around zero is not rejected up to 10% level. So the inclusion of longer lags to create normality on error terms is having more predominant impact on reducing the power of the test than improving the performance of the test through white noise ness of the error terms.

In general the logical conclusion is that all prices are following unit root but the existence of two unit roots, which is less common in economics is highly improbable in the wheat prices of Ethiopia. But it is rational to keep reservation for price of Addis Ababa, Mekelle and Nazareth since there is slim probability that they can be stationary. However the search procedure used in this paper will buffer them out in process of searching for rule of one price, if they are stationary. So there is no logical reason to exclude them in the cointegration analysis. This is so since the evidence toward unit root is much stronger than the contrary evidence. Given these facts now let's focus on the cointegration analysis given next.

3.2. <u>Cointegration analysis under one common trend</u>

Numbers of markets are assumed to be cointegrated, if they are experiencing flow of goods and information from one to another to develop long run relationship between the prices discovered in different markets. But strong form of cointegration will be observed if two or more markets are not only experiencing flow of goods and information but are also operating under rule of one price means they are having one common trend. If markets are under rule of one price any intervention in any of the markets will be automatically felt on other markets, even though some markets may have better influence than others. These facts can be analyzed by using Johansen (1998 and 1991) vector error correction model and the related methodology developed Gonzalo and Granger (1995) to estimate the single common trend and its determinants. However these methodologies assume that there is no persistent trade reversal in estimating the same coefficients for all periods. If X is supplying Y in 6 months and Y is supply X in the next 6 months, the cointegration coefficients will be deferent in each half of the year. There is need to account for such change by using switching regression which is not done in Johansen vector error correction model and related models. So these models are appropriate only and only if there is no persistent trade reversal between markets.

If there is reliable flow data it can be used to analyze the possibility of trade reversal but there is no reliable flow data for Ethiopia; so data on price reversals are used to analyze the possibility of trade reversal. Since the focus is in persistence trade reversal which can have affect on the long run relationship between market prices, average monthly prices

than series of monthly prices are used. And the average monthly prices used for the analysis are given in figure 5, below.

**Figure 5 Average monthly white wheat prices in 8 white wheat markets of Ethiopia (1996-2003 G.C.)**
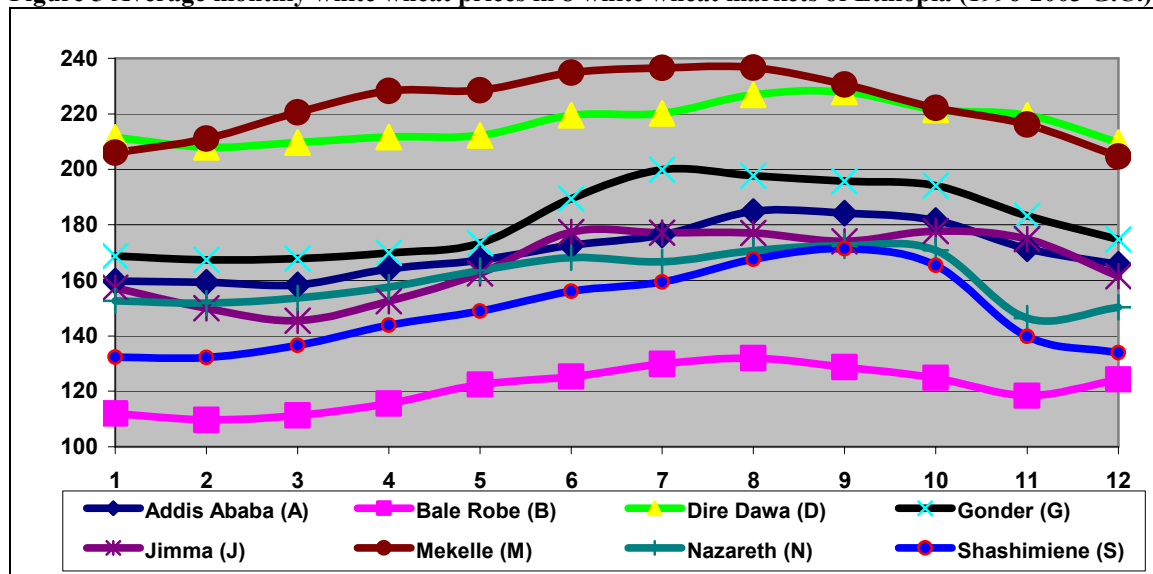
The highest average prices are observed in two deficit markets of Mekelle at North and Dire Dawa at East (see figure 6, below). Given these markets are found in considerable distance from each other the observed crossing in their average monthly price can't be due to trade reversal. Following Mekelle and Dire Dawa the highest price is observed in Gonder at North West side of the county which is not having any price reversal with any of the markets. The lowest monthly price is observed in surplus market of Bale Robe at South followed by another surplus market of Shashimiene in South central. And both markets are not showing any evidence of price reversal with any of the markets. So for the five markets which include three deficit markets of Mekelle, Gonder and Dire Dawa and two surplus markets of Bale Robe and Shashimiene trade reversal is not a series problem.

However the deficit market of Jimma at West is having persistence price reversal in some months of the year with both central market of Addis Ababa and Secondarily central market of Nazareth. So if Jimma is found to be un-cointegrated with either Addis Ababa or Nazareth, it is not logical to conclude that there is no relationship between these markets; since it can be related to trade reversal. In general excluding Jimma all other markets can be effectively analyzed by using Johansen vector error correction model since there is no persistence trade reversal. However the analysis below will show that Jimma is highly cointegrated with 5 out 7 markets used in the study. So trade reversal is not series issue in these paper.

The next issue is related to the question: which one of the 5 models of Johansen vector error correction models have to be used? Theoretically the most logical models are three. Vector error correction model with restricted constant with assumption that the long run average log of transaction cost is constant, with out trend in level data. Vector error

44

correction model with unrestricted constant can be used to allow for existence of deterministic trend in the level data in addition to constant log of transaction cost. But figure 1 in page 37 did clearly show that there is no visible trend in the level data. But to allow for use of central seasonal dummy with out allowing any trend in log of transaction cost, this model can be also be very useful.

**Figure 6 Ethiopian Map and Markets under this study**
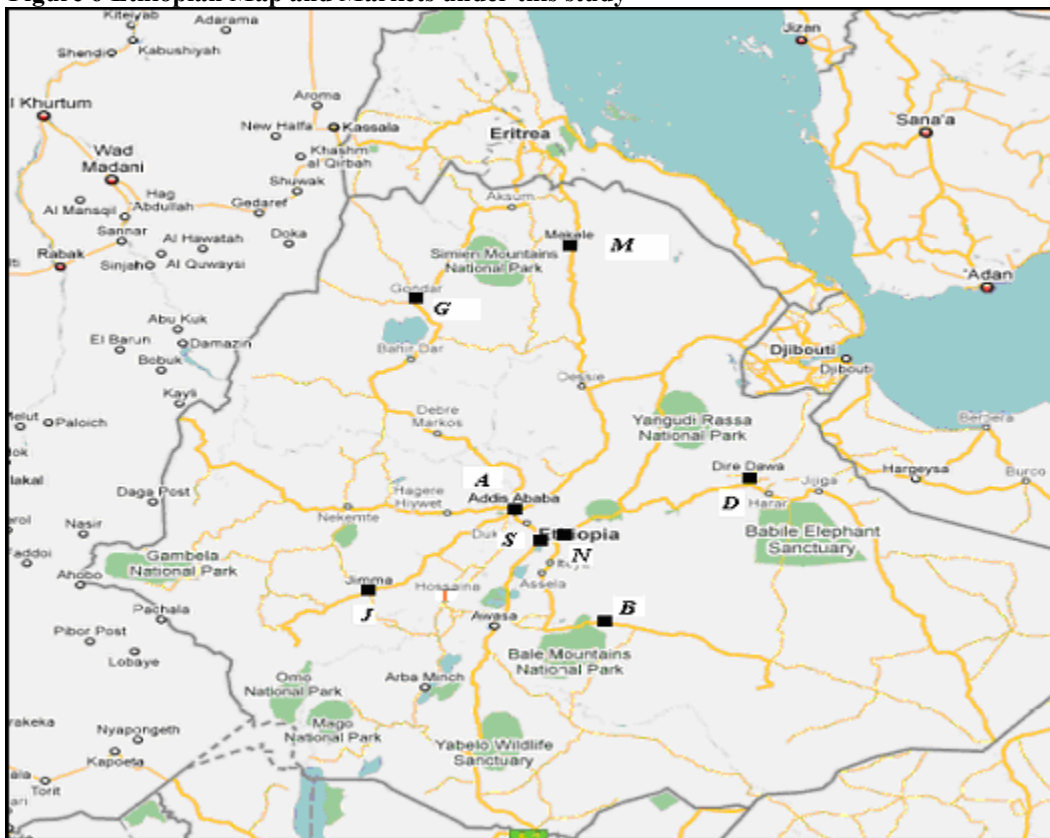
Source 3 Google map (with miner modification)

The third one is vector error correction with restricted trend. This will allow for trend on log of transaction cost. In this paper these three models are considered as possible option, but not the two other models. Vector error correction with out any deterministic variable is illogical since it assumes zero log of transaction cost. And vector error correction with unrestricted trend assumes quadratic trend on level data but the data used like most economic data is not showing deterministic quadratic trend.

But to determine if restricted trend or restricted constant are more appropriate some blurred evidence can be observed from log of price margin between markets. And for simplicity and convenience the log of price margin between Addis Ababa which is central hub and the rest of the markets is given in figure 7, below. And as can be seen from the figure most of the log of price margins are not trended except for some form of declining random trend in price differential between Mekelle and Addis Ababa and to some extent between Addis Ababa and Dire Dawa. So for most market combinations the right model seems to be error correction model with restricted constant. But the problem

with this model is it is not possible to control the impact of seasonality through centered seasonal dummy which demands the use of unrestricted constant in the model. For that purpose unrestricted constant model is also fitted for all market permutations.

Figure 7 Price difference between Addis Ababa and 7 major wheat markets
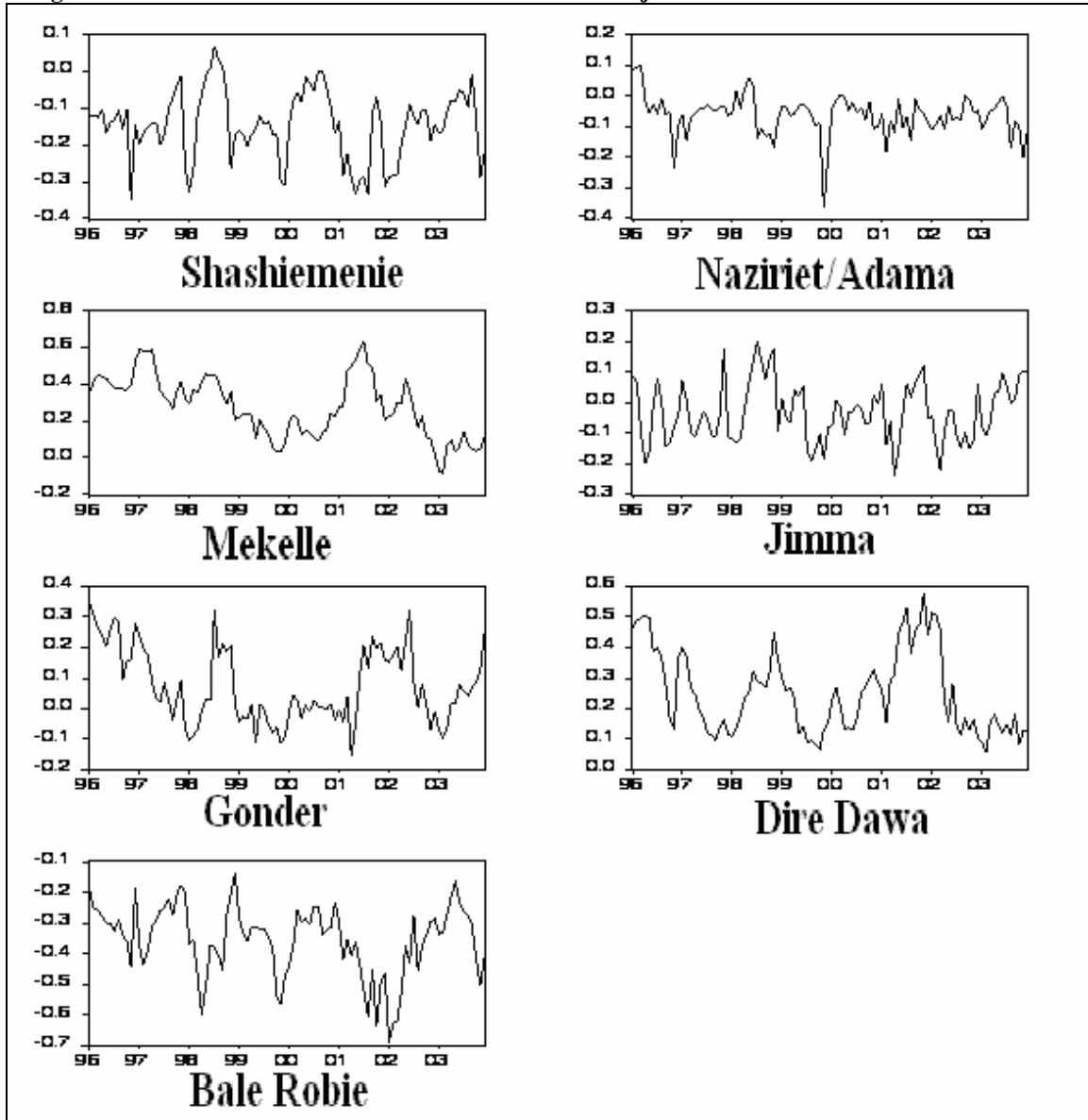
The third model of restricted trend is also used to allow for possibility of trend in log of transaction cost of some markets like Mekelle. Log of transaction cost is not observed and what is available is log of price margin. Moreover log of price margin is log of transaction cost and profit. So it is not logical to assume for all markets there is no trend in their log of transaction cost given there is no trend in log of price margin. Since log of price margin is blurred indicator of log of transaction cost dynamics, restricted trend model is also fitted for all market permutations. It is left to the data to select the appropriate model for each combination of markets from the three models.

3.3. <u>The basic idea of markets ruled by one price</u>

The basic idea of cointegration and common trend can be explained if we think about drank person's walk. Each next step in walk of drank person is random addition from his current position. He can have deterministic trend with random impulse, if he is going some where, say his house or next bar. Or if he is not going any where but simply hustling around he may only have random variation or random trend around his initial place. Now again think about his dog and his wife, who are following him. Following his pattern their walk will also have random impulse at each step. In vector error correction model it will be tried to estimate the long run relationship between the drank-man, his dog and his wife. There will be two long run relationships, one between the drank-man and his wife and another relationship between the drank-man and his dog. But all of them are following one common trend which is generated by the drank-man. If we know their common trend we can know their long term random pattern. And understanding who is influencing the direction, in this case the drank-man we can focus our intervention in just one person. Since the drank-person is the main determinant of the common trend and the dog and the wife are going where ever the drank is going, we can influence all of them by targeting the drank-person. Johnson (1998 and 2001) vector error correction model will be help full in determining the long run relationship between variables and Gonzalo and Granger (1995) model will be used to identify the long run common trend and the major determinants of the common trend.

But if the drank have another friend who is walking to ward his own destination with them, because he is in the same lane. There will be long run relationship between the friend and all others, but you can't fully control his moment by controlling his drank friend. So you will have two long run relationships and two common trends. One relationship will be between the first drank-man and his friend, and the second relationship will be between the first drank man his wife and his dog.

In this case you will have two common trends. One is generated by the drank-man and another is generated by the friend who is walking to his own destination. The friend is walking with them because they are in the same way, so you can't completely influence his moment by controlling his drank friend. Strong intervention in one side can even break the cointegration all in all. In this case you need to influence both friends at the same time for effective out come. So identification of common trends and their major determinant is very critical for effective intervention with low administration cost.

Applying the analogy to wheat markets in spatially dispersed markets the search for one common trend between cointegrated markets is the search for markets under rule of one price. If markets are governed by rule of one price intervention in one market will be felt in every market. But most importantly if the market/markets which is/are determining the common trend is/are identified intervention can be very effective, if it is targeting this/these market/markets.

In this paper a routine search is done starting from two dimensions up to possible dimensions, for all possible lag orders, to identify markets which are sharing one

common trend. Lag selected by information criterions is not used to choose lag order. But for those market permutations which are following one common trend given the null hypothesis of normality and no serial correlation can't be rejected up to 10% level, information criterions are used to see if the fitted lag is in line with lag order selected by different information criterions.

One problem is that of deciding which information criterions to use when they are in conflict with each other. Most Monte Carlo Studies did not consistently show any one measure to be superior to any other measure. To give some example Khim−Sen, (2004) focusing in one dimension fourth order auto regressive regression with stationary series found that all criterions i.e. Hannan and Quinn information criterion (HQIC), Final prediction error (FPE), Schwartz's Bayesian information criterion (SBIC) and Akaike's information criterion (AIC) are picking the right lag in more than 50% of the cases. In small samples (less than 120) AIC and FPE are having the lowest probability of under fitting compared to HQIC and SBIC. And over fitting is not a series problem for all criterions. So based on this study AIC and FBE seems to be preferable to others.

But the above analysis did not consider non stationary series and higher dimensional VAR systems. Gutiérrez, Souza and Guillén (2007) taking cointegrated two dimensional vector error correction model found that for samples size of 100, which is close to sample size used in this paper, AIC is performing better in picking the right lag order compared to HQIC and SBIC. And most often all are selecting the right rank of one. One additional advantage of AIC is related to the fact that AIC is found to be insensitive to restrictions putted in adjustment parameters. This is additional evidence that AIC is better in small samples even though FPE is not taken in to account in this study.

However Chao and Phillips (1999) by considering 3 dimensional vector error correction fitted by Johansen method with out any deterministic or moving average component found very contradictory result. In small samples AIC is found to be inferior to SBIC. This is so if performance measured by appropriate lag selection or identification of lag which resulted in to right identification of rank. And the bias is not observed to decline with sample size means the criterions is found to be not only biased but also inconsistent.

But a more general Monte Carlo simulation with different data generating process by Gonzalo and Pitarakis (2000) found that the relative performance of each criterion is sensitive not only to data generating process assumed in relation to dimension, unit root, moving average component but also to the maximum lag allowed in selection process. This is clearly putted in their statement

> *'These mixed and often contradictory conclusions clearly highlight the point that it is difficult to come up with a universally accepted typology of methods ranked in terms of their performance. Indeed the number of factors influencing the behavior of these procedures is such that conclusions can only be DGP specific, with different parameterizations possibly leading to contradictory features for the same criterion.' (page 2)*

Mathematical proof and Monte Carlo simulation given by Gonzalo and Pitarakis (2000) shows that AIC theoretically and practically can have higher probability of over fitting and it is inconsistent in large samples. But the bias of under fitting well decline to zero as sample size goes to infinitive means AIC is consistent for bias of under fitting. This is the case for small dimension of 1 and 2. But for higher dimension more than or equal to 3 the bias will disappear due to dimensional impact. As result AIC is also found to be not only consistent on higher dimensions but also the over fitting bias is small in large samples. And this is the case for stationary, none stationary and cointegrated vectors and series.

The study also shows that even though on low dimensions HQIC and SBIC are having low probability to over fit and are consistent. They are also observed to under fit and rarely observed to move from lower lags on high dimension vectors causing gross under fitting.

But given moderately small sample size used in this paper, let's focus on simulated result under small sample size. In small sample with low dimension the same study found that AIC will have inflated Chi Square distribution resulting on high level over parameterization. But it will select appropriate lags in samples size as moderate as 90. But if the dimension is increased to more than 2 AIC will only have high probability of over fitting, if the sample size is very small compared to number of parameters that have to be estimated. Unfortunately, if the number parameters that have to be estimated are large in number AIC have high probability of over fitting.

In general the study show that SBIC to be the preferable in terms of it lower probability of over fitting in small and large samples. But the probability under fitting seems to be series in case of SBIC. On large samples under fitting is improbable in all models but in small samples SBIC will normally have high probability of under fitting but AIC have negligible probability of under fitting problem. But if the system dimension increases all criterions have higher probability to under fit the right model. However when the true lag order and the dimension is higher the under fitting problem of AIC is lower than HQIC and SBIC, which are rarely observed to move from the lower bound.

But their relative performance in terms of over fitting is observed to be dependent not only on sample size and degree of freedom in the estimation, but also on the maximum lag allowed in information criterions. In small sample size as moderate as 100 AIC is observed to have lower probability of over fitting compared to others. But if the maximum lag allowed is higher AIC is observed to pick the maximum lag allowed in most cases. In general on high dimensional and small sample size AIC seems to be better than other measures. Moreover if there is moving average component on the data which needs longer lag to be approximated by auto regressive terms to attain white noise error vector AIC is more robust than all information criterions. However a study by Ng and Perron (2001) found that, if there are roots close to -1 in moving average component both AIC and SBIC have tendency to under fit the true model.

In general one measure did not seem to be better than others in all cases as performance is observed to be dependent on unobserved data generation processes existing in the data.

49

But this does not mean that we can pick any lag as right lag. A study by Winker and Maringer (2004) shows that not only the relative performance of the criterions is data generation process dependent but also unlike classical regression analysis over fitting is as worst problem as under fitting. Study by Ho and Sørensen (1996) as sited in Maringer (2004) also indicate that when longer lags are used in appropriately, Johansen's rank test have tendency to over estimate the number of cointegration equations found in the model.

So because the Monte Carol evidence is conflicting, it does not mean we can pick any lag order as right lag order, since it can result in erroneous conclusion. Moreover the maximum lag possible is assumed to be known but it is not. And the Monte Carlo studies, stated above, did show that the selection of maximum lag allowed can have series impact on performance of the information criterions. In this paper a pragmatic approach is used to deal with this problem. First for 2 and 3 dimensions a minimum of maximum lag of 10 is used and this is increased up to maximum possible. And in each stage the selected lag by all information criterions is recorded. And the modal selected lag is reported as selected lag. On higher dimensions given low degree of freedom the possible maximum lag is small, so in order to have better representative modal lag, it was found necessary to reduce the minimum of maximum lag allowed. As result for 4th and higher dimensions the minimum of maximum lag allowed is reduced from 10 to 5. Given these methodological facts the cointegration analysis is given below.

3.4. Cointegrated markets under rule of one price in bivariate dimension
3.4.1. For model with restricted constant

If large number of markets are under rule of one price each of them are pair wise cointegrated. However the two dimensional estimation will biased compared to the *n* dimensional estimation (Gonz´alez-Rivera and Helfand 2001). So error vectors generated from biased coefficients may fail to have white noise distribution. But most importantly the rank test which is based on data generating process with normally, independently and identically distributed error vectors, for moderately small sample size used in this paper, will not be accurate. So the rank test may not find the markets to be cointegrated with one common trend in lower dimensions and even if they are identified to be cointegrated with one common trend, their error vector may not be white noise. If the error vector is not white noise you can't take the rank test result as sound justification for cointegration since it has higher tendency to find false cointegration. So the search procedure followed by Gonz´alez-Rivera and Helfand (2001) and Rashid (2006) is not theoretically sound but let's see if the problem is series in real empirical data.

The search is started from all possible permutations of two markets. For model with restricted constant out of 1680 permutations[12] only 16 permutations are found to have one rank given the assumption of normally and independently distributed error terms and vectors can't be rejected up to 10% level for each market and lags, respectively.

---

[12] For 8 markets at each lag there are $8 \times 7 = 48$ permutations and in 30 lags there are $8 \times 7 \times 30 = 1680$ permutations

Table 5 for model with restricted constant rank test in two dimensions

Markets / critical value	Lags	Trace statistics		Maximum Eigen value	
		0 rank	1 rank	0 rank	1 rank
J – A	1	27.4	2.4	25	2.4
J – A	3	29.77	6.78	22.99	6.78
J – A	22	30.11	5.11	25	5.11
A – B	23	29.71	3.16	26.55	3.16
A – B	24	29.46	3.77	25.68	3.77
M – A	26	47.34	5.87	41.47	5.87
J – G	30	78.4	7.18	71.22	7.18
B – D	24	38.28	3.47	34.81	3.47
S – N	24	50.75	2.89	47.86	2.89
A – S	14	34.44	9.17	25.27	9.17
A – S	15	29.76	7.24	22.52	7.24
5%		19.96	9.42	15.67	9.24
1%		24.60	12.97	20.20	12.97

And as can be seen from table 5 above for all market permutations given above the null of 0 rank is rejected at 1% level, but the null of only one cointegration equation can't be even rejected at 5% critical level by both testes. So the market permutations at their specified lag are par wise cointegrated with one common trend. But Johnson rank test is based on the assumption of correctly specified error correction model in terms of lag order and distribution of the error vectors. But most importantly it is very sensitive for slight misspecification of lag order (Chao and Phillips 1995) and distribution of the error vector mainly serial correlation (Hamilton 1994 and Johansen 1998 and 2001). So specification test are given in table 6, below.

Table 6 for model with restricted constant lag, normality & serial correlation testes for two dimensions

Markets	Lag used	Lag selection by different information criterions				Jarque - Bera test for Normality		LM - Serial correlation test	
		FPE	AIC	HQIC	SBIC	Statistic (4 – df)	Prob	Statistic (60 –df)	Prob
J – A	1	3	3	2	2	1.986	0.738	61.499	0.422
J – A	3	3	3	2	2	1.77	0.778	61.362	0.427
J – A	22	3	3	2	2	1.645	0.801	45.811	0.912
A – B	23	1	1	1	1	2.353	0.671	42.129	0.961
A – B	24	1	1	1	1	3.114	0.539	34.274	0.997
M – A	26	3	2	2	2	1.608	0.807	56.868	0.591
J – G	30	2	2	2	1	1.874	0.759	61.164	0.434
B – D	24	2	5	1	1	5.11	0.276	50.715	0.798
S – N	24	2	2	2	2	3.154	0.532	53.458	0.422
A – S	14	2	2	2	2	2.066	0.724	51.524	0.422
A – S	15	2	2	2	2	2.177	0.703	58.32	0.422

Normality is observed to be sensitive to order of markets and this is expectable given the fact that in reduced rank regression if Y = AX + e is true, then X = (1/A)Y + e/A is not

necessarily true (Hamilton 1994). In this paper given sample size of 96 observations strict normality is demand in each error term of each market and error vector of the system. In all cases unless the null of zero skewness, zero kurtosis and normality can't be rejected up to 10% level the combination is defined as unfit. But to save space only the vector level Jarque - Bera (J-B) test of normality is given. And among all permutations of given combination the one with the highest probability of normality is given in table 5 and 6 above. As can be seen from table 6 above the null of normality can't be rejected up to 27% level for all market permutations.

For serial correlation LM test is used and the test is carried up to 15th lag given monthly nature of the data. Again strict independence of error vectors is demanded in each serial lag from 1 to 15. Unless the null of independence can't be rejected up to 10% level at each lag, the combination is dropped as unfit. However again in order to save space only vector level results are given. Since the distribution of the error vector under the null are following Chi square distribution and they are independent of each other, their sum will also follow chi Square distribution with degree freedom equal to their total degree of freedom. And as can be seen from the table 6 above all permutations of markets are having error vector which are not only normally but also independently distributed.

To observe if the lag fitted is backed by information criterions the selected lag and fitted lag are also given in table 6, above. Four kinds of information criterions are used in this study; these are Akaike's information criterion (AIC), Final prediction error (FPE), Hannan and Quinn information criterion (HQIC) and Schwartz's Bayesian information criterion (SBIC). As was stated before Monte Carlo estimations by Gonzalo and Pitarakis (2000) did show that unless great care is made in selection maximum lag AIC have tendency to select the maximum allowed lag. And in this study as maximum allowed lag approaches to ward maximum in above case 30, not only AIC but also some times HQIC and rarely FPE and SBIC are observed to jump to ward the maximum lag allowed. To avoid this problem for each market combination lags are selected 21 times by using maximum lags in range of 10 to 30. And the selected lag will be the modal lag. Since there is no scientific formula for selecting the maximum lag and the testes are sensitive to the maximum lag allowed, this is a pragmatic solution used in these paper to deal with the problem.

And if strict normality and independence conditions are imposed, only Jimma and Addis Ababa are well cointegrated markets at bivariate level or with in two dimension assuming constant log of transaction cost and un-trended level data. Other markets are having highly over fitted lag order and it is more probable that their observed cointegration is related to the fact that in small samples and over fitted longer lags the rank test is known to having very weak power to reject false null hypothesis. So the central market of Addis Ababa which is observed to handle 35% of the national trade by studies done as early as 1995/6 (Gebremeskel et al, 1998) and Jimma a deficit market which is also a major market for coffee originated in western part of the country is found to be well cointegrated with the central market. First Jimma is not only wheat deficit market but also located in major coffee supplying market, which is the main export item of the country. Means there is real demand backed by purchasing power. Second most of the

coffee whole sellers can use their experience, network and capital developed in coffee trading in creating effective space utility in imperfect grain market. And this will make grain traders from coffee supplying markets well capitalized, well connected and well informed with adequate marketing experience to have better performance in creating efficient space utility. This will make Jimma a major trading center following Addis Ababa. So the strong cointegration is found between capital city and major zonal market in the country. This is in line with the finding of Gonz´alez-Rivera and Helfand (2001) and Rashid (2006).

However if strict normality demanded for each error term is dropped and only vector level normality is demanded and if the demand for strict independence at each serial lag is replaced by demand for independence only at vector level, in addition to cointegration between Addis Ababa and Jimma at two lags, additional two cointegrations are observed to make Bale Robe, Jimma and Addis Ababa pair wise cointegrated to each other. As can be seen from table 7 below, in the three combinations there is only one common trend in each combination. This is so because the null of 0 rank is rejected with 99% confidence by both trace statistics and Maximum Eigen value, but the null of one common trend can't be rejected at 5% level.

Table 7 for model with restricted constant rank test in two dimensions (with imperfect specification)

Markets/ critical values	Lags	Trace statistics		Maximum Eigen value	
		0 rank	1 rank	0 rank	1 rank
A - B	1	28.4196	2.6163	25.8033	2.6163
J - A	2	30.5394	5.6580	24.8814	5.6580
B - J	1	25.7230	2.4781	23.2449	2.4781
5%		19.96	9.42	15.67	9.24
1%		24.60	12.97	20.20	12.97

In the above analysis it is observed that Addis Ababa and Jimma are observed to have strong cointegration at 1st and 3rd lag. But information criterions are picking 2nd and 3rd lags and cointegration is observed in 1st and 2nd lag. So this clear evidence that the deficit market of Jimma and central market of Addis Ababa are well cointegrated. In second lag the null of normality is not rejected at 57% and the null of independence on error vector is not rejected at 10% level in any cumulative serial lag between 1st and 15th (see table 8 and 9, below).

Table 8 for model with restricted constant lag, normality and serial correlation testes for two dimensions (with imperfect specification)

Markets	Lag used	Lag selection by different information criterions				Jarque - Bera test for Normality		LM - Serial correlation test	
		FPE	AIC	HQIC	SBIC	Statistic (4 – df)	Prob	Statistic (60 –df)	Prob
A - B	1	1	1	1	1	6.910	0.141	39.74	0.98
J - A	2	3	3	2	2	2.918	0.572	72.487	0.129
B - J	1	2	2	2	2	2.229	0.694	57.264	0.576

Moreover at first lag which is backed by all information criterions cointegration is observed between Addis Ababa and Surplus market of Bale Robe. For this combination the null of normality is not rejected at 14% level and the null of independent error vector can't be rejected at 51% level at all cumulative serial lags from 1st to 15th.

Third cointegration is between Bale Robe and Jimma at one lag. All information criterions are picking second lag as right fit but given their possible up ward and down ward bias these criterions are not taken as perfect predictors of the lag order than as educated guess work to ward the true lag order. So it is not rejected as wrong fit. For this combination the null of independent vector is not rejected at 17% level at all possible cumulative serial lags and the null of normality of the error vector is not rejected at 69% level.

Table 9 for model with restricted constant the highest serial correlation for 2 markets (with imperfect specification)

Markets	LM statistics	Serial cumulative lag	Degree of freedom	Probability
A-B	19.044	5	20	0.519
J-A	69.691	14	56	0.103
B-J	6.398	1	4	0.171

One final test is related to existence of ARCH effect on the data. ARCH effect is observed to have insignificant impact on asymptotic distribution of rank test. See Johansen (1988 and 1991), Cavaliere et al (2009) and Hamilton (1994) for theoretical exposition and Lee and Tse (1996) and Cavaliere et al (2009) for Monte Carlo evidence. However in small samples rank test is dependent on assumption of identically and independently distributed Gaussian error vectors. And Monte Carlo evidence by Lee and Tse (1996) did show that when there is ARCH effect rank test have tendency to over reject the null of no-cointegration and the bias will increase with the size of the ARCH parameters mainly the first parameter of the first squared difference. For example when there is no heteroskedasticity the power of trace test with 100 observations, which is close to the observations used in the study, is 98.56% and the relative value for Maximum Eigen value is 99.06%. However if there is ARCH effect with the first parameter equal to 0.1 and second parameter equal to 0.8 the power will decline marginally to 97.97% and 98.34%, respectively. However if the first parameter is increased to 0.3 and the second reduced to 0.6 the power will decline to 95.94% and 96.37%, respectively. In general the impact of ARCH is to reduce the power of the rank test defined in terms of rejecting false cointegration but the impact is very marginal. Means there is low probability that cointegration can be found when there is none but the probability depends on the size of the parameters and mainly the first parameters. Armed with these facts lets analyze the result below.

For the 3 combinations which are showing independently distributed error vector with single common trend ARCH test is done and is given in table 10 below. For cointegration of Jimma and Addis Ababa from 1 to 3 lag and Addis Ababa and Bale Robe in 1st lag the null of constant variance is not rejected at 14% level. So for these markets there is no problem heteroskedasticity. However the cointegration observed between Bale Robe and Jimma is also observed to have time varying heteroskedasticity in Bale Robe error term.

So given the probability of marginally lower power of rank test it is logical to doubt the cointegration between these two markets.

Table 10 ARCH effect test for model with restricted constant at 2 dimensions

Combination	Lags	markets	Wald test		
			Wald statistics	Degree of Freedom	Probability
J - A	1	J	11.5	8	0.174
		A	2.1	1	0.14
J - A	2	J	0.21	1	0.6431
		A	2.36	2	0.3072
J - A	3	J	0.31	1	0.5798
		A	1.48	2	0.4761
A- B	1	A	0.6	1	0.412
		B	0.07	1	0.794
B – J	1	B	47.3	8	0.000
		J	0.7	1	0.386

However the first parameter is found to be very small with value of -0.0055861 and the over all sum of the eight coefficients is also only 0.2100066. To avoid negative coefficient in ARCH regression, it is re-estimated   by imposing the conditions needed to make variance always positive in GivWin soft ware with GARCH version 1 programming. And the value of first coefficient turns out to be numerically and statistically close to zero. And the sum of eight coefficients is 0.220036331 and mainly caused by third coefficient in ARCH regression. Under these restrictions ARCH effect for Bale Robe is rejected at 0.57% level. So what ever ARCH effect there is not that strong to introduce large bias in the power of the rank test. So the combination is accepted as following one random trend.

In general assuming constant log of transaction cost and level data with out any deterministic trend the three markets are pair wise cointegrated. The deficit market of Jimma at west is showing the strongest cointegration with central market of Addis Ababa. The second stronger cointegration is observed between the central market of Addis Ababa and Surplus market of Bale Robe at South. And the third relatively weakest cointegration, if strength is measured by distributional character of the error vector, is observed between Bale Robe and Jimma. This could be an indirect indicator that Addis Ababa is serving as central market for both markets in line with early study by Gebremeskel et al (1998) and Asfaw (1998).

3.4.2. **For model with unrestricted constant**

Even though there is no visible trend in prices of all grains unrestricted constant is used in order to allow the use of seasonal dummy to control seasonal effect on monthly price of grains. Centered seasonal dummies are introduced to control monthly and quarterly seasonality in price. One quarterly dummy is used for months of June to August when major agriculture production related activities are performed. Second dummy is used for months of September to November in which most of the harvesting is done. For

December to January when the harvest is marketed in large quantity the third dummy is used and one last dummy for the rest of the months. And for each month one dummy is used to account monthly seasonality. Given the above fact let's analyze the result next.

Table 11 for model with unrestricted constant rank test in bivariate dimension

Markets/ critical value	Lags	Trace statistics		Maximum Eigen value	
		0 rank	1 rank	0 rank	1 rank
A – M	22	21.666	2.252	19.414	2.252
A – M	27	20.932	1.623	19.31	1.623
D – B	24	37.218	0.147	37.07	0.147
B – S	24	27.058	0.038	27.02	0.038
J – D	23	26.662	2.681	23.981	2.681
N – D	21	24.588	0.517	24.071	0.517
A – B	1	24.021	1.998	22.024	1.998
J – A	1	24.716	1.819	22.897	1.819
D – M	24	60.267	0.29	59.978	0.29
A – S	15	27.885	3.654	24.231	3.654
5%		15.41	3.76	14.07	3.76
1%		20.04	6.65	18.63	6.65

As can be seen from the table 11, above all market permutations are having one rank. In all cases the null of no cointegration is rejected at 1% level but the null of 1 rank is not rejected at 5% level. So assuming the lag length is right and the model with unrestricted constant is right model all the combinations given in table 11 are having one common trend.

Table 12 for model with unrestricted constant lag, normality and serial correlation testes for 2 markets

Markets	Lag used	Lag selection by different information criterions				Jarque - Bera test for Normality		LM - Serial correlation test	
		FPE	AIC	HQIC	SBIC	Statistic (4 – df)	Prob	Statistic (60 –df)	Prob
A – M	22	11	11	2	1	3.247	0.517	49.398	0.834
A – M	27	11	11	2	1	1.819	0.769	48.868	0.847
D – B	24	5	5	1	1	1.899	0.754	55.42	0.644
B – S	24	2	2	2	1	2.535	0.638	48.765	0.85
J – D	23	7	8	1	1	1.888	0.756	49.576	0.829
N – D	21	1	1	1	1	1.187	0.88	53.094	0.724
A – B	1	2	2	1	1	5.806	0.214	38.863	0.984
J – A	1	3	3	3	1	1.523	0.823	60.506	0.457
D – M	24	1	1	1	1	5.129	0.274	42.399	0.834
A – S	15	2	2	2	1	1.875	0.759	57.019	0.834

As can be seen from table 12 above for all market permutations the assumption of normality is not rejected and the error vectors are independently distributed in all cases. So based on normality and serial correlation all permutations correctly specified. The

problem is related to the fact that there is significance difference between fitted lag and lag selected by information criterions for most markets. But as was stated before some information criterions mainly AIC, and to some extent others are sensitive to the maximum lag allowed in the information criterions. Just if allowed maximum lag is increased by small margin, say from 20 to 21, the selected lag is observed to jump dramatically, say from 2 to 21. In this paper maximum lags are selected in the interval of 10 to 30 and the modal lag selected it taken as right lag order.

Again even though 9 different combinations are observed to have strongly cointegrated white noise error vector the lag order fitted by 7 combinations of markets is not supported by any of the information criterions. The exceptions are the cointegration between central market of Addis Ababa and one surplus market of Bale Robe and one deficit market of Jimma. In one direction the surplus market of Bale Robe is well cointegrated with Addis Ababa and in other direction the deficit market of Jimma is well cointegrated with Addis Ababa. The combination of Addis Ababa and Bale Robe is using1 lag and this is supported by both HQIC and BSIC but rejected by AIC and FPE. However the difference is one lag and given poor performance of AIC on low dimension and small sample it is accepted as right fit. The cointegration of Jimma and Addis Ababa is having lag of 1 which is only supported by SBIC, only. According to the three other criterions it is under fitted by two lags. Again there is strong evidence to reject the cointegration between these markets, so they are accepted as having one common trend.

Table 13 Likelihood test for unrestricted constant two markets

Combination	With out seasonal indicators		With seasonal indicators	
	LM statistics (1)	Prob.	LM statistics (22)	Prob.
J - A	0.0011	0.96	46.5171	0.002
B - A	0.019	0.891	48.308	0.001

The important question is which one of the two models i.e. vector error correction with restricted or unrestricted constant is the parsimonious model. 2 times the difference between log likelihood of unrestricted constant and restricted constant has chi square distribution with $n - h = 1$ degree of freedom or number of common trends (Johansen 1988 and 1991). And as can be seen above the null of restricted constant with out any trend in level data can't be rejected up to 96% level for Addis Ababa and Jimma, and 89% level for Addis Ababa and Bale Robe. However when seasonal dummy are used the seasonal dummy is significant at 1% level for both. Means even though there is no trend in level data, inclusion of unrestricted constant is justifiable given the need to control for seasonal effect. Out of 11 monthly dummies 3 are dropped due to multicollinearity with the three quarterly dummies. As result there are 11 seasonal dummies in each equation total of 22 dummies in both equations. So the test on seasonal indicators has 22 degree of freedom. In the first test the null is restricted constant and the alternative is unrestricted constant with out seasonal dummy. In second test the null is unrestricted constant with out seasonal dummy and the alternative is with seasonal dummy. In general, which ever model is selected i.e. restricted or unrestricted constant model, the central market of Addis Ababa is found to be cointegrated with major deficit and major surplus market and

again this is in line with the finding of Gonz´alez-Rivera and Helfand (2001) and Rashid (2006).

When serial correlation and normality condition are dropped 11 different combinations are observed to have one common trend under lag order which is supported by some of the information criterions. However even though serial correlation was not series problem normality was not attained even at 1% level. The distribution of most error vectors except one was far from normal. And given the fact that the rank test has tendency to find wrong cointegration with none normal errors in moderately small sample, they are rejected as unfit.

Table 14 for model with unrestricted constant lag, normality and serial correlation testes for two markets (with imperfect specification)

Markets	Lag used	Lag selection by different information criterions				Jarque - Bera test for Normality		LM - Serial correlation test	
		FPE	AIC	HQIC	SBIC	Statistic (4 – df)	Prob	Statistic (60 –df)	Prob
D – B	1	5	5	1	1	7.598	0.107	53.675	0.705

So with out demanding normality for each error term and independent error vector for each serial lag, but only white noise error vector for all markets in the combination, the combination of Dire Dawa and Bale Robe is observed to have one common trend. Moreover its lags are supported by HQIC and SBIC but not by AIC and FPE. However given in small samples and low dimensions AIC have tendency to over fit, the combination is accepted as right fit. The null of normality at vector level can't be rejected at 10% level and independence of the error terms based on test of cumulative serial lag of 15 is also not rejected up to 70% level. But since each serial lag is not insignificant the number of lag selected can have impact on conclusion reached. For that purpose for all possible serial lags from 1 to 15, cumulative LM test is done. And the lowest probability is observed at 5th serial lag with LM statistics of 28.7779. Given in 5th lag the degree of freedom is equal to 20 the relative probability is equal to 9.2%. Means the white noise nature of the Dire Dawa and Bale Robe cointegrated error vector is not very strong. This is so because information criterions are also picking 5th serial lag as right lag. In over all However the serial correlation is not series enough to result on rejection of the cointegration, since it is insignificant at 5% level.

Table 15 for model with unrestricted constant rank test (with imperfect specification)

Markets/ critical values	Lags	Trace statistics		Maximum Eigen value	
		0 rank	1 rank	0 rank	1 rank
D-B	1	20.3525	3.7354	16.6172	3.7354
5%		15.41	3.76	14.07	3.76
1%		20.04	6.65	18.63	6.65

Observing the rank test given in table 15 the cointegration of Dire Dawa and Bale Robe at 1st lag is supported by trace statistics which is rejecting rank of 0 with 99% confidence but is not rejecting the null of one common trend at 5% level. But the maximum Eigen value is having contradictory result in which at 1% level it is finding 0 cointegration

between these markets but at 5% there is one cointegration. So again this is additional evidence that the cointegration between Dire Dawa and Bale Robe is very weak, if it exists at all.

The final test for ARCH effect is given in table 16 below. For both cointegration of Jimma and Addis Ababa and the cointegration of Bale Robe and Addis Ababa the null of constant variance is not rejected at 64% level. So it is logical to conclude that there is no heteroskedasticity problem for both combinations. However in the cointegration of Dire Dawa and Bale Robe heteroskedasticity is observed in the Dire Dawa error term with 99% significance. Means it is questionable if there is cointegration between these markets since the rank test has tendency to over find cointegrated vectors with existence of ARCH effect on error term.

Table 16 ARCH effect test for model with unrestricted constant at 2 dimensions

Combination	Lags	markets	Wald test		
			Wald statistics	Degree of Freedom	Probability
J - A	1	J	0.07	1	0.7954
		A	0.21	1	0.6479
B- A	1	B	0.04	1	0.8400
		A	0.21	1	0.6460
D- B	1	D	49.18	5	0.0000
		B	0.13	1	0.7167

However again the first coefficient is only 0.0638609 and the total sum of 5 coefficients is 0.4808595 which implies the impact of ARCH effect on power of the testes is very marginal. So it is logical to accept that the strongest cointegration is found between Addis Ababa and one deficit market of Jimma and surplus market of Bale Robe in 2 dimensions. Additionally, there is possibly weak cointegration between Dire Dawa and Bale Robe.

3.4.3. For model with restricted trend

The restricted trend model is useful if log of transaction cost has deterministic trend. Under inflationary situation, continuous devaluation or increase in economic activity the log of transaction cost may show up ward trend over time. However it does not seem to be the case in most grain markets observed from 1996 to 2003. In the same period inflation was very low and even negative in some years[13], the foreign exchange rate though determined by auction between banks still it is not fully liberalized. Moreover the State has high role in stabilizing the value of foreign exchange. Oil price is again stabilized by government buffer fund so the domestic economy is shielded as much as possible from international oil price volatility. But the improvement on the quality of road in some part of the country could have impact on generating negative trend on log of transaction cost. This is the case for Mekelle and to some extended for Dire Dawa, which are having negative stochastic trend on their log of price margin from Addis Ababa.

---

[13] To given example in 2000 inflation was 6.2% but in 2001 it was negative 5.2 and in 2003 negative 7.2 (AFDB, 2007)

For others the reduction on log of transaction cost due to improvement on transportation facility could be counter balanced by the observed meager inflation and devaluation. The most important transaction costs are related to transportation. And with in transportation repair and fuel are the most important cost components. The improvement on road quality will reduce the cost of spare parts and inflation and devaluation will increase the fuel and spare part cost. And the data on price margin indirectly shows that there is no visible trend on it. Means it is more probable that they are balancing each other out to eliminate any possible trend. However since we lack data on log of transaction cost we can't rule out the use of restricted trend model for any of the markets. It is left for the data to define if restricted trade is appropriate for some of the markets or not.

Table 17 for model with restricted trend rank test for two markets

Markets/ critical value	Lags	Trace statistics		Trace critical values at 1%	
		0 rank	1 rank	0 rank	1 rank
A – B	21	35.368	4.744	30.623	4.744
A – B	22	47.236	9.949	37.287	9.949
J – A	22	41.927	11.648	30.279	11.648
M –A	5	40.33	7.375	32.954	7.375
A – M	7	38.147	7.286	30.86	7.286
A – M	15	32.263	6.49	25.774	6.49
B – J	12	38.418	7.708	30.71	7.708
J – D	17	33.345	8.355	24.989	8.355
J – D	24	32.334	10.064	22.27	10.064
D – J	25	50.706	10.144	40.563	10.144
J – M	24	54.109	8.237	45.872	8.237
M – G	25	60.985	5.949	55.035	5.949
A – S	14	36.08	7.488	28.592	7.488
5%		25.32	12.25	18.96	12.52
1%		30.623	16.26	23.65	16.26

As can be seen from table 17 above out of 1680 permutations of markets tested only 24 permutations are having 1 rank with normally and independently distributed error vectors. Assuming that the markets are properly specified, the rank test shows that all bivariate market combinations are having one common trend. This is so because the null of zero cointegration is rejected for all markets at 1% level but the null of one rank or one cointegration equation can't be rejected even at 5% level. But the rank test is based on the assumption white noise error vector and appropriate lag selection. If the model is well specified or not is tested and given in table 18 below.

All combinations are having strongly normal error vector since the null of zero skewness and kurtosis can't be rejected at 10% level for each error term and the null of vector level normality is not rejected even at 42%, for all market combinations. Moreover the null of independently distributed error vector is not rejected at each serial lag from 1st to 15th and at the cumulative serial lag of 15there is no evidence of serial correlation even at 37% level for all markets.

Table 18 for model with restricted trend lag, normality and serial correlation testes for two markets

Markets	Lag used	Lag selection by different information criterions				Jarque - Bera test for Normality		LM - Serial correlation test	
		FPE	AIC	HQIC	SBIC	Statistic (4 – df)	Prob	Statistic (60 –df)	Prob
A – B	21	2	4	1	1	1.636	0.802	46.695	0.895
A – B	22	2	4	1	1	1.428	0.839	49.644	0.827
J – A	22	3	5	3	2	3.192	0.526	40.42	0.975
M –A	5	5	5	5	1	2.703	0.609	46.521	0.899
A – M	7	5	5	5	1	2.836	0.586	48.877	0.847
A – M	15	5	5	5	1	3	0.558	42.768	0.955
B – J	12	8	9	2	1	2.172	0.704	48.074	0.866
J – D	24	8	8	1	1	3.487	0.48	48.074	0.866
J – D	25	8	8	1	1	1.146	0.887	45.995	0.909
D – J	17	8	8	1	1	0.073	0.999	48.013	0.868
J – M	24	5	5	1	1	1.771	0.778	45.775	0.912
M – G	25	5	5	1	1	3.836	0.429	62.849	0.376
A – S	14	2	2	2	1	2.601	0.627	56.228	0.614

The problem comes from the lag used to fit combination. Studies did show the fact that in moderately small samples like the sample used in this paper and if large lags are used that rank test have weak power to rejected false null hypothesis. And as can be seen from information criterions except the cointegration between Mekelle and Addis Ababa, all other combinations are observed to have one common trend because they are over fitted. The combination of Addis Ababa and Mekelle have cointegration with white noise error vector at 5, 7 and 15 lags; and three of the information criterions except SBIC are picking lag of 5 as correct lag. So at five lags the deficit market of Mekelle which is located in Northern drought porn area of the country is cointegrated with Addis Ababa. Moreover the declining common trend observed in case of Mekelle log of price margin from Addis Ababa is caused by declining log of transaction cost of creating space utility.

Since the log likelihood test demand that the null and alternative hypothesis to be properly specified and the combination of Mekelle and Addis Ababa are not cointegrated at other models, it is not possible to test if the trend in log of transaction cost is statistically significant or not. The next question is if the strict normality and serial correlation condition imposed in the model are relaxed which combination of markets can be found to be cointegrated.

When strict independence, normality and lag conditions are dropped many market combinations are observed to have one common trend. But given the fact that rank test is developed under assumption of independently distributed error vector for all sample sizes and normally distributed error vector for small sample sizes, it is not logical to accept the rank test result on market combinations with out white noise error vectors. Moreover rank taste is also dependent on the assumption of right lag selection. Fortunately, the level of serial correlation and normality problem is to be found less series in some market combinations given in table 19, below. Moreover there is no evidence to reject their lag

order as their lag is supported by some of the information criterions and there is no conclusive evidence about relative performance of each information criterions.

Table 19 for model with restricted trend lag, normality and serial correlation testes for two dimensions (with imperfect specification)

Markets	Lag used	Lag selection by different information criterions				Jarque - Bera test for Normality		LM - Serial correlation test	
		FPE	AIC	HQIC	SBIC	Statistic $(4 - df)$	Prob	Statistic $(60 - df)$	Prob
A – G	17	5	17	2	1	6.606	0.1582	55.656	0.635
J – B	9	8	9	2	1	2.604	0.6262	55.82	0.629
J – M	5	5	5	1	1	4.367	0.3586	54.854	0.664
M – B	8	9	9	1	1	1.616	0.806	68.468	0.212

The four market combinations given above are free from series none normality problem. However there is some evidence of serial correlation. Serial correlation is observed in some serial lags but the over all serial correlation in all 15 serial lags is insignificant up to 21% for all. But we can't conclude that there is no serial correlation problem, because serial correlation is observed when lower lags are selected. The exception is the cointegration between the surplus market of Bale Robe and Deficit market of Jimma in which the highest level of serial correlation observed is at 11th lag with LM statistics of 48.7. However with 44 degree of freedom it is only significant at 28.9% level. Moreover the lag order used is supported by AIC and missed by one lag by FPE. However HQIC and SBIC are selecting lower lags and the combination is grossly over fitted according to both criterions. In small samples and in low dimensions AIC has high probability over fitting but others have high probability of under fitting. So we can't completely rejected or accept the market combination as right or wrong fit with acceptable confidence. However given the fact that FPE is also picking 8th lag, it is impossible to reject the market combination as unfit. So it is logical to conclude there is cointegration between Jimma and Bale Robe at 9 lag assuming trended log of transaction cost. The same is observed in case of restricted constant model in which Bale Robe and Jimma are observed to have one common trend at first lag. So the restricted trend model is nesting the restricted constant model by unrestricting the constant and by adding one restricted trend and 8 additional lags in each equation or 16 total lags in both. The effect can be tested by using log likelihood test in which 2 times the difference on their log likelihood will have Chi square distribution with 18 degree of freedom. The log likelihood for restricted trend at 8th lag is 240.0611 and the relative figure for restricted constant at 1st lag is 189.9429; as result 2 times their difference is 100.24, which is significant at 1% level. So even though it can't be sure, if the model or the lags are contributing to significance of the general model, the restricted trend with higher lags is having better fit compared to restricted model of restricted constant. And since it was not possible to find one rank at 1st lag it was not possible to separate the impact of the lags and the unrestricted trend. Just for curiosity restricted trend model is fitted at 1st lag imposing cointegration and it is compared against restricted constant. The LM statistics becomes 58.78, which is significant at 1% level with 2 degrees of freedom. So the possibility of trend in log transaction of cost of Jimma and Bale Robe is statistically significant.

Again Mekelle seems to be cointegrated with both Bale Robe and Jimma. In 15th cumulative serial lags there is no series serial correlation problem in these combinations. However in some serial lags evidence of serial correlations are observed. Taking Mekelle and Jimma the highest serial correlation is observed at 3 cumulative serial lag with LM value of 20.4906 which is significant at 10% but not 5% level. And this serial lag is selected by information criterions. For Bale Robe and Mekelle the highest serial correlation is observed is at first serial lag with LM value of 12.847 which is significant at 5% but not at 1% level. Again this is also backed by information criterions. So the cointegration of Mekelle and Bale Robe is very questionable but the cointegration of Mekelle and Jimma is having low level of serial correlation, so it is not logical to reject it.

Table 20 for model with restricted trend the highest serial correlation and serial correlation on lags selected by information criterions for two markets (with imperfect specification)

Markets	LM statistics	Serial lags	Degree of freedom	Probability
A – G	20.174	2	8	0.009697
J – B	48.7025	11	44	0.289447
J – M	20.4906	3	12	0.058356
M – B	12.8479	1	4	0.012043
A – G	7.1663	1	4	0.12736
J – M	20.4906	3	12	0.058356
M – B	12.8479	1	4	0.012043

So the deficit market of Mekelle at north seems to be more cointegrated deficit market of Jimma in South west than the surplus market of Bale Robe at South. Another cointegration observed is between deficit market of Gonder in North West and central market of Addis Ababa at 17th lag. But this combination have two problems, first the 17th lag used is only backed by AIC, which is know to have up ward bias, but not by others which are selecting very low lag. And normally rank test have weak power when higher lags are used inappropriately. Second at 2nd serial lag the highest over all serial correlation is observed which is significant at 1% level. Moreover the maximum Eigen value is defining the combination as un-cointegrated but not trace statistics (see table 21, below). So the observed cointegration between Gonder and Addis Ababa is very questionable, too.

Table 21 for model with restricted trend rank test (with imperfect specification)

Markets/ critical values	Lags	Trace statistics		Maximum Eigen value	
		0 rank	1 rank	0 rank	1 rank
G – A	17	33.8239	12.1299	21.6940	12.1299
B – J	9	35.3642	8.2642	27.1000	8.2642
M – J	5	32.6209	4.2005	28.4204	4.2005
B – M	8	46.3966	11.5084	34.8882	11.5084
5%		25.32	12.25	18.96	12.52
1%		30.45	16.26	23.65	16.26

And the fact that it is having independent error vector at first serial lag, which is selected by information criterions, can' change the above conclusion, though it can't imply that the serial correlation problem is less series.

As can be seen from table 21, above, for pair wise combinations of Jimma, Bale Robe and Mekelle the null of 0 rank is rejected at 1% level but the null of one common trend can't be rejected even at 5% level. So it is logical to conclude that the three markets are pair wise cointegrated with each other and Addis Ababa, having the strongest cointegration with central market of Addis Ababa under different models (assumptions).

ARCH effect is tasted for all combinations and is given in table 22 below. For all combinations, except one, the null of constant variance can't be rejected at 22% level. However the cointegration between Addis Ababa and Gonder is having significant ARCH effect related to Gonder at 10% level but not at 5% level. However again the only coefficient in variance equation is equal to 0.1204911, which is very small to introduce significant bias in to the rank test. So if the cointegration between Addis Ababa and Gonder has to be rejected, it can't be due to heteroskedasticity but other factors stated above. So based on ARCH test there is no reason to reject the conclusion that all market combinations are having one common trend and are ruled by rule of one price.

Table 22 ARCH effect test for model with restricted trend at 2 dimensions

Combination	Lags	markets	Wald test		
			Wald statistics	Degree of Freedom	Probability
M –A	5	M	0.20	1	0.6519
		A	0.11	1	0.7352
A – G	17	A	0.00	1	0.9806
		G	2.73	1	0.0983
J – B	9	J	0.00	1	0.9515
		B	1.48	1	0.2244
J – M	5	J	0.04	1	0.8507
		M	0.17	1	0.6826
M – B	8	M	0.01	1	0.9279
		B	0.02	1	0.8963

Before triple markets are analyzed some important points about normality test in vector correlation needs to be pointed out. Even though Jarque – Bera test for normality can be used for both vector error correction and vector auto regressive models with out any modification (Demiroglus, 2000); it is found to be order sensitive in this paper. This is in line with the fact that the choice of normalization can have effect on coefficient estimated in vector error correction model unless the variables are perfectly correlated (Hamilton, 1994). Moreover in higher dimensions the order of the variables used in addition to choice of normalization of the variables can have impact on value of coefficients estimated. This is due to nature of the canonical correlation used in Johnson vector error

correction model. In above bivariate case the combination with the highest normality is given from the two possible permutations of two markets. If the markets are given as X – Y, it means the coefficient of X is normalized to one, and this permutation is with highest normality.

In higher dimensions say three, if X – Y – Z are given as market permutation it means. First the strongest cointegration is between X and Z, given X's coefficient is normalized to be one. And the second cointegration is between Y and Z independently of the relationship of X and Z, in which the coefficient of Y is normalized to be one. And in this paper permeation of markets with highest possible normality from all possible permutation are given in all dimensions.

3.5.   <u>Cointegrated markets under rule of one price in triplet dimension</u>

Now it is time to add a third market which shares the same long run trend with the two markets. In Gonz´alez-Rivera and Helfand (2001) study of Brazil rice market recursive testes are done to search for the rule of one price. And they found that the formation of the boundary of rule of one price to be insensitive to order of inclusion. But actually the reason for that conclusion is because they were not testing for normality at each stage, which is order sensitive. They found that distance is the main important variable in determining order of inclusion. And in Rashid (2006) study of Uganda maize market distance is used to determine the order of inclusion taking the finding of Gonz´alez-Rivera and Helfand (2001) in to consideration. In this paper recursive test are done for all possible orders and at each stage rank, serial correlation, normality and ARCH testes are made to make sure the combination is having one common trend with white noise error vector. More ever it is checked if the lag used is in line with the lag selected by the four information criterions.

The permutation of markets is accepted as right fit only and only if it has one common trend and if the assumption of normally and independently distributed error variance can't be rejected up to 10% level of significance. In serial correlation test it is demanded that for each serial lag from 1 up to 15 lags, each serial lag's coefficient have to be insignificant up to 10% level. For normality in addition to vector level normality for each market's error term the null of zero skewness and kurtosis must not be rejected up to 10% level of significance. This needed a barrage of tests on large permutations of markets. For restricted constant 7392 permutations are fitted and tested. The relative figure for both restricted trend and unrestricted constant is 6720 permutations[14]. Leaving aside it's computational and time cost disadvantage, the advantage of this methodology is that any unproven assumption will not be imposed in the search procedure and all assumptions are left to be proven from the data. One of the important assumptions of earlier application of this methodology is that there is no strength reversal which is not assumed in this paper, for example. If 3 or more markets are following one common trend all bivariate combinations are also cointegrated with one rank. However bivariate estimation will result on biased coefficients (Gonz´alez-Rivera and Helfand 2001). So it is possible the

---

[14] The difference is resulted from the fact that only 20 lags can be fitted for the last two but it is possible to fit 22 lag for the first.

error terms in bivariate combination can not be white noise. And if the error terms are not white noise rank test is not perfectly reliable in small samples. So it is not necessary that all higher dimension cointegrations have to be built, on lower dimension combination of cointegrated markets found in the higher dimension cointegration. Means there is a probability that strength reversal can be observed in higher dimensions. In this paper contrary to Gonz´alez-Rivera and Helfand (2001) and Rashid (2006) recursive routine testes are done with out ignoring the possibility of strength reversal. And the triplet dimension results are given below starting from restricted constant model.

3.5.1. <u>For model with restricted constant</u>

Out of 7392 permutations only 34 permutations are having three markets with one common trend and normally and independently distributed error vectors. The 34 permutations are divided in to 12 market combinations. In each combination the permutation with the highest probability of normality is given in table 23, below.   And as can seen from table 23 below for all combinations, except one, the null of 0 and 1 rank are rejected with 99% confidence by both tastes. However the null of one common trend can't be rejected at 5% level by both testes. So assuming there is no specification problem it is logical to conclude that the 11 combinations are having one common trend.

Table 23 for model with restricted constant rank test for three markets

Markets/ critical value	Lags	Trace statistics			Maximum Eigen value		
		0 rank	1 rank	2 rank	0 rank	1 rank	2 rank
J–D–A	17	81.58	26.72	1.78	54.86	24.94	1.78
B–G–A	17	70.24	34.17	5.56	36.06	28.62	5.56
J–A–B	1	53.82	27.33	2.28	26.48	25.06	2.28
J–D–B	18	62.15	31.71	2.04	30.44	29.67	2.04
G–A–D	17	63	25.62	0.68	37.38	24.93	0.68
M–D–J	17	56.35	26.81	8.59	29.53	18.23	8.59
N–J–D	20	95.31	41.64	6.6	53.67	35.03	6.6
N–J–D	21	120.76	37.16	8.74	83.61	28.41	8.74
S–N–D	13	58.1	24.76	1.83	33.34	22.92	1.83
M–A–J	18	110.95	48.33	5.73	62.62	42.6	5.73
M–A–J	19	107.14	47.25	7.58	59.89	39.67	7.58
S–G–J	17	62.36	31.49	6.8	30.87	24.69	6.8
5%	■	34.91	19.96	9.42	22.00	15.67	9.24
1%	■	41.07	24.60	12.97	26.81	20.20	12.97

The exceptional market is that the cointegration between Jimma, Addis Ababa and Bale Robe. They are having one common trend based on trace statistics but the conclusion of Maximum Eigen value is mixed. Based on 5% critical level of maximum eigen value they are having one common trend. But at 1% level the null of 0 rank can't be rejected but the null of 1 rank is rejected. This is inconsistency in conclusion of Maximum Eigen value since they can't have 2 ranks with out having 1 rank. So these markets are accepted as cointegrated markets under rule of one price based on rank test, not only because the evidence to contrary is inconclusive but also because they are found to be strongly

cointegrated at lower dimension. The logical conclusion is that all 12 combinations are having one rank, assuming there is no specification problem.

As was stated before serial correlation is tasted for each serial lag and a permutation is assumed as right fit only and only if the null of no serial correlation is not rejected for each serial lag from 1 up to 15 lags at 10% level. But for presentation convenience the over all LM test is given in which the statistics is simple sum of each lag's statistics and as was stated before this have Chi square distribution with degree of freedom equal to sum of the individual lags degree of freedom ($15 \times 9 = 135$ in this case). As can be seen from table 24 below, in all market combinations there is no serial correlation problem. This is so because the null of independently distributed error vectors can't be rejected up to 28% level of significance for all market combinations. Moreover the null of normality as tested by Jarque – Bera test can't be rejected up to 16% level for all market combinations. Additionally for each error term in any of the permutations skewness and kurtosis test is done and in all cases the null of zero skewness and kurtosis is not rejected at 10% level. So all combinations are having independently distribution Gaussian error terms and vectors as demanded by rank test.

Table 24 for model with restricted constant lag, normality and serial correlation testes for three markets

Markets	Lag used	Lag selection by different information criterions				Jarque - Bera test for Normality		LM – Serial correlation test	
		FPE	AIC	HQIC	SBIC	Statistic (6– df)	Prob	Statistic (135-df)	Prob
J–D–A	17	2	2	1	1	8.203	0.224	143.963	0.283
B–G–A	17	2	2	2	1	0.285	1	116.807	0.869
J–A–B	1	2	2	2	1	4.934	0.552	121.454	0.792
J–D–B	18	2	2	2	1	1.271	0.973	135.511	0.471
G–A–D	17	5	5	1	1	1.137	0.98	121.666	0.788
M–D–J	17	2	2	1	1	1.763	0.94	126.294	0.692
N–J–D	20	1	1	1	1	5.164	0.523	104.369	0.976
N–J–D	21	1	1	1	1	9.084	0.169	120.765	0.805
S–N–D	13	3	3	2	1	7.117	0.31	113.059	0.915
M–A–J	18	2	2	2	1	3.844	0.698	122.323	0.775
M–A–J	19	2	2	1	1	4.398	0.623	103.719	0.283
S–G–J	17	2	2	2	1	3.422	0.754	119.885	0.283

However except the combination of Jimma, Bale Robe and Addis Ababa all other combinations are found to be over fitted based on all information criterions. So given the fact that when the model is over fitted on small samples rank test have tendency to find cointegration when there is not any, it is logical to conclude other markets are not triple wise cointegrated with each other, with one common trend.

The deficit market of Jimma which is also an important source of the main export commodity of the country, coffee is having common trend with surplus market of Bale Robe and central market of Addis Ababa. The lag is justified by SBIC but is found to be

under fitted based on other three information criterions. Monte Carlo studies stated above did show that any single information criterion is not superior over others in terms of performance. AIC has tendency to over fit and HQIC and SBIC have tendency to under fit. But in higher dimensions and sample size around 100, it is also shown that AIC is not inferior compared to others in terms of its probability to over fit. But still the probability of under fitting will be higher among all testes in high dimension. Given this facts it seems the combination of Jimma, Addis Ababa and Bale Robe is under fitted by one lag. However given the bivariate cointegration observed between these markets and given the fact that SBIC is backing it, it accepted as good fit. And this is logical given the fact that the probability of success by most testes found in most Monte Carlo studies is always around 50% to 70% not close to 100% in any condition. In this paper the information criterions are used to test if the specification is grossly deviating from the selected lags, not as accurate selectors of the appropriate lag.

ARCH effect is tested for the cointegration observed between Jimma, Addis Ababa and Bale Robe and the null of constant variance is rejected with 99% confidence for Bale Robe (see table 25, below). However the first coefficient in variance equation of Bale Robe is just -0.004778 and the sum of the eight coefficients is just 0.2035776 so this is not big enough to introduce strong bias on the rank test. Since the coefficient can't be negative in order to avoid negative conditional variance the needed restrictions are imposed in parameters. And the first coefficient turn out to be as small as 0.0958279 with total sum of 0.396261309 for the eight parameters. And after using efficient maximum likelihood estimation test for ARCH effect is done and the statistics turn out to be equal to 3.79025 which is only significant at 87.55% level only. As result these three markets are accepted as following one common trend.

Table 25 ARCH effect test for model with restricted constant at 3 dimensions

Combination	Lags	markets	Wald test		
			Wald statistics	Degree of Freedom	Probability
J–A–B	1	J	0.53	1	0.4654
		A	0.62	1	0.4306
		B	45.01	8	0.0000

By relaxing the condition on normality from term level to vector level and serial correlation from serial lag level to cumulative serial lag level it was not possible to find additional three market combinations with one common trend. From this result two things can be observed first as expected by Gonz´alez-Rivera and Helfand (2001) the cointegration is found to be built on markets which are cointegrated at lower dimension. However distance does not seem to be an important factor since Bale Robe in south is located in considerable road distance from Jimma in west and Addis Ababa in center. Bale Robe and Jimma are spatially close to each other, However they area connected only by road that passes through Addis Ababa. Since the search of Gonz´alez-Rivera and Helfand (2001) is started form *m* major markets, if the search is started from combination of these three markets it will be in line with their study. But if the analysis is started from two markets and others are added based on their distance from capital city as Rashid (2006) neither Jimma nor Bale Robe will be the first choice. Shashimiene and

Nazareth are much closer to Addis Ababa than either Jimma or Bale Robe. So even though it has to proven farther in higher dimensions distance did not seem to be the only important factor for order of inclusion in to one price system.

3.5.2. <u>For model with unrestricted constant</u>

In unrestricted constant model out of 6720 permutations of markets 43 different permutations are found to have one common trend with white noise error vectors. These are divided in to 11 combinations given in table 26, below. All combinations of markets, except 2, are having one common trend as both the null of 0 and 1 rank are rejected at 1% level, by both testes but the null of 2 ranks is accepted at 5% level. If there is no specification problem it is logical to conclude that the 9 combination of markets are following one common trend.

Table 26 for model with unrestricted constant rank test for three markets

Markets/ variables	Lags	Trace statistics			Maximum Eigen value		
		0 rank	1 rank	2 rank	0 rank	1 rank	2 rank
G–D–A	17	61.5	26.21	0.52	35.29	25.68	0.52
M–J–A	18	134.84	54.85	2.98	79.99	51.86	2.98
M–J–A	19	185.94	69.32	3.39	116.62	65.93	3.39
J–A–N	17	87.62	31.79	2.67	55.83	29.12	2.67
J–A–N	18	80.77	25.64	1.04	55.13	24.6	1.04
J–A–B	1	46.47	21.99	1.71	24.48	20.27	1.71
J–D–B	7	40.23	20.1	2.62	20.14	17.48	2.62
J–D–B	18	60.73	21.59	1.36	39.13	20.23	1.36
N–D–B	17	89.54	36.89	0.44	52.66	36.45	0.44
N–D–B	18	79.86	34.04	0.07	45.82	33.96	0.07
S–J–D	14	48.51	22.94	1.07	25.57	21.87	1.07
N–G–J	20	116.88	53.43	2.2	63.45	51.23	2.2
5% critical value		29.68	15.41	3.76	20.97	14.07	3.76
1% critical value		35.65	20.04	6.65	25.52	18.63	6.65

Jimma, Addis Ababa and Bale Robe again are having one common trend based trace statistics at both 1% and 5%, level but not based on Maximum eigen value. As before at 5% level maximum Eigen value is founding one common trend, but not at 1% level. At 1% the maximum Eigen value is inconsistent as the null of 0 cointegration equations is not rejected but the null of only 1 cointegration equation is rejected. Given the markets are pair wise cointegrated and maximum Eigen value is inconsistent in its conclusion, they are accepted as following rule of one price.

The combination of Jimma, Dire Dawa and Bale Robe is having one common trend based trace statistics at both 1% and 5% level. However based on maximum Eigen value the null of 0 rank is not rejected at 1% and 5% level and the rank of 1 is rejected at 1%, but not at 5% level. So there is no conclusive evidence to accept or reject the rank test between these three markets.

 As can be seen from table 27 below the null of independently distributed error vector is not rejected at 49% level in all combinations; so the error vectors of all combinations are found to be independently distributed. Again the permutation of markets given in table 26 and 27 are of those permutations of markets with highest level of normality. And as can be seen in table 27 below the null of normality can't be rejected up to 70% level for all combinations. So the error vectors are normally and independently distributed. But the question is: are they having acceptable lag as defined by the information criterions?

Table 27 for model with unrestricted constant lag, normality and serial correlation testes for three markets

Markets	Lag used	Lag selection by different information criterions				Jarque - Bera test for Normality		LM - Serial correlation test	
		FPE	AIC	HQIC	SBIC	Statistic (6– df)	Prob	Statistic (135-df)	Prob
G–D–A	17	5	13	1	1	3.108	0.795	117.802	0.854
M–J–A	18	3	10	1	1	3.819	0.701	117.678	0.856
M–J–A	19	3	10	1	1	2.902	0.821	105.728	0.97
J–A–N	17	5	5	1	1	3.221	0.781	113.009	0.916
J–A–N	18	5	5	1	1	1.716	0.944	106.249	0.968
J–A–B	1	5	5	1	1	3.752	0.71	119.256	0.831
J–D–B	18	5	7	1	1	2.478	0.871	109.14	0.95
J–D–B	7	5	7	1	1	2.119	0.908	134.507	0.496
N–D–B	17	1	1	1	1	2.483	0.870	134.006	0.508
N–D–B	18	1	1	1	1	6.175	0.404	131.967	0.558
S–J–D	14	1	1	1	1	3.679	0.72	131.727	0.564
N–G–J	20	9	12	1	1	2.573	0.86	109.313	0.949

Except for two combinations of three markets the rest combination of three markets are grossly over fitted and their cointegration is not supported by any of the information criterion. So it is logical to conclude that their apparent one common trend is result of weak power of rank test on over fitted loner lags than cointegration of the prices under one common trend. The cointegration of Jimma, Bale Robe and Addis Ababa at one lag is supported by both HQIC and SBIC. However FPE and AIC are picking 5th lag. Again the information criterions are in conflict with each other. On higher dimensions AIC have better performance but given the markets are pair wise cointegrated and the cointegration of the three markets is not high dimension enough to give more emphasis on AIC, the market combinations at one lag are accepted as good fit.

The second combination is between surplus market of Bale Robe and deficit markets of Jimma at South West and Dire Dawa at east at 7th lag. FPE is picking 7th lag defining it right fit, AIC is picking 5th lag defining it to be over fit by two lags. However based on HQIC and SBIC it is grossly over fitted by 6 lags. So the evidence is not conclusive. One thing is clear that if we expect all information criterions to pick one lag order, it is not possible to pick any lag as right one. So even though it can not be granted the combination Jimma, Dire Dawa and Bale Robe at 7th lag to be the right fit, there is no conclusive evidence to reject it, either.

Moreover the weak cointegration observed between surplus market of Bale Robe at South and deficit market of Dire Dawa at East is also built in to another relatively weaker cointegration under rule of one price by adding deficit market of Jimma at west if 1 lag is used and strict distributional assumptions are relaxed. And as can be seen from table 28, 29 and 30 below if strict normality and serial correlation condition are relaxed these markets are having one common trend which is backed by trace statistics with white noise error vector and lower lag of one.

Table 28 for model with unrestricted constant rank test for three markets (with imperfect specification)

Markets/ critical value	Lags	Trace statistics			Maximum Eigen value		
		0 rank	1 rank	2 rank	0 rank	1 rank	2 rank
JDB	1	41.079	21.093	3.287	19.987	17.806	3.287
5%		29.68	15.41	3.76	20.97	14.07	3.76
1%		35.65	20.04	6.65	25.52	18.63	6.65

However Maximum Eigen value is rejecting any cointegration at 1% level and at 5% level the conclusion is mixed. Again there is no logically conclusive evidence to reject or accept the markets as following one common trend. But accepting they are following one random trend the used lag is backed by HQIC and SBIC but not others. And the null of vector level normality is not rejected at 37% level and serial correlation is not rejected at 15% level for any cumulative serial lag from 1 to 15 (See table 29 and 30 below).

Table 29 for model with unrestricted constant lag, normality and serial correlation testes for three markets (with imperfect specification)

Markets	Lag used	Lag selection by different information criterions				Jarque - Bera test for Normality		LM - Serial correlation test	
		FPE	AIC	HQIC	SBIC	Statistic (6– df)	Prob	Statistic (135-df)	Prob
JDB	1	5	7	1	1	6.416	0.378	129.382	0.62

So the 7th lag used to find one common trend under strict distributional assumption is needed to make the error terms more Gaussian and independent compared to first lag. However in both lags there is high probability that there is one common trend between these markets. Now let's observe ARCH effect below.

Table 30 for model with unrestricted constant the highest serial correlation for three markets (with imperfect specification)

Markets	LM statistics	Serial lags	Degree of freedom	Probability
JDB	13.2358	1	9	0.152224

Observing ARCH effect given in table 31 below it can be observed that inclusion of unrestricted constant not only can allow the use of centered seasonal dummy but also it is observed to eliminate the ARCH effect for combination of Addis Ababa, Bale Robe and

Jimma. For Jimma, Dire Dawa and Bale Robe at 7th lag the null of constant variance is not rejected at 73%, but at 1st lag the null is rejected in case of Dire Dawa with 99% confidence. However the first coefficient is negative and small with value of -.0702418. Imposing restrictions on the parameters to make the conditional variance always positive the Best unbiased estimator based on maximum likelihood test is used and the first coefficient turn out to be as small as 0.0182917 with total sum of the 5 parameters as small as 0.018378938. So it is not logical to reject the existence of one common trend based on ARCH effect not only at 7th lag but also at 1st lag. However the use of 1st lag, except to make the lag used consistent with some information criterions selected lag, did not seem to add any thing to the performance of the market as the rank test is still inconclusive and some miner ARCH effect is introduced.

Table 31 ARCH effect test for model with unrestricted constant at 3 dimensions

Combination	Lags	markets	Wald test		
			Wald statistics	Degree of Freedom	Probability
J–A–B	1	J	0.33	1	0.5678
		A	0.11	1	0.7455
		B	0.05	1	0.8147
J–D–B	1	J	0.17	1	0.6839
		D	55.06	5	0
		B	0.49	1	0.4837
J–D–B	7	J	0.04	1	0.8375
		D	0.12	1	0.7302
		B	0.06	1	0.8063

The cointegration of Jimma, Addis Ababa and Bale Robe even assuming trended level data is found to be built on those markets which are cointegrated in two dimensional levels. And the importance of unrestricted constant can be tested by using log likelihood test as was stated before. When seasonal indicators are dropped the test will follow Chi square distribution with $n - h$ degree of freedom or one in this case. But when 11 seasonal indicators are used in each equation making total of 33 seasonal indicators in 3 equations plus one degree of freedom for unrestricted constant will result on 34 degree of freedom.

Table 32 Likelihood test for unrestricted constant three markets

	With out seasonal indicators		With seasonal indicators	
Model	LM statistics (1)	Prob.	LM statistics (34)	Prob.
J–A–B	0.144	0.704336	71.9284	0.0001567

And as can be seen from table 32 above, the model with restricted constant is more fit than the model with unrestricted constant. However the seasonal indicators are very significant to warrant the use of unrestricted constant to allow for seasonal effect. Moreover given the elimination of ARCH effect which can improve the performance of rank test it is preferable to use unrestricted constant than restricted constant model.

In general there is strong cointegration under rule of one price between Bale Robe, Jimma and Addis Ababa but the rule of one price between Jimma, Dire Dawa and Bale Robe is questionable if unrestricted constant model is used.

3.5.3. **For model with restricted trend**

Out of 6720 permutations of markets with restricted trend only 83 permutations are found to be having one single common trend with normally and independently distributed error vectors. And these are made of 19 combinations of markets and these combinations are given in table 33 below. Out of the 19 combinations, in which some are having replication with different lag, trace and maximum Eigen value are finding one common trend in 16 combinations. So if there is no specification problem these markets are cointegrated under rule of one price. The exceptions are, one the cointegration between Dire Dawa, Shashimiene and Nazareth in which even though trace statistics is founding one common trend, maximum Eigen value is inconsistent in its conclusion at 1% level.

Table 33 for model with restricted trend rank test for three markets

Markets/ critical value	Lags	Trace statistics			Maximum Eigen value		
		0 rank	1 rank	2 rank	0 rank	1 rank	2 rank
D–M–A	17	85.03	43.79	7.09	41.24	36.69	7.09
G–J–A	18	142.06	46.57	10.12	95.49	36.46	10.12
B–A–S	16	81.97	35.25	11.44	46.71	23.81	11.44
B–D–S	12	71.52	36.54	7.74	34.99	28.8	7.74
D–S–B	18	121.37	41.46	8.3	79.92	33.16	8.3
D–J–G	11	65.31	32.03	5.28	33.27	26.75	5.28
D–J–G	19	148.65	47.5	7.86	101.15	39.64	7.86
M–D–J	12	63.4	30.82	6.53	32.58	24.29	6.53
S–J–D	18	74.93	39.12	10.74	35.8	28.38	10.74
D–S–N	14	61.43	31.62	6.35	29.81	25.27	6.35
S–N–D	17	121.57	47.84	3.56	73.72	44.28	3.56
D–S–N	18	132.05	37.35	4.92	94.7	32.43	4.92
J–B–G	20	540.16	192.44	7.47	347.72	184.98	7.47
J–M–G	12	96.02	42.56	8.32	53.46	34.24	8.32
D-N-G	17	74.93	31.52	9.19	43.42	22.33	9.19
J–N–G	18	95.29	44.47	9.32	50.82	35.15	9.32
N–J–M	14	71.35	39.51	11.45	31.85	28.05	11.45
B–N–M	15	76.09	32.19	8.19	43.91	23.99	8.19
B–N–M	18	149.65	34.61	12.07	115.05	22.53	12.07
5%		42.44	25.32	12.25	25.54	18.96	12.52
1%		48.45	30.45	16.26	30.34	23.65	16.26

The second one is cointegration of Dire Dawa, Nazareth and Gonder in which trace statistics is founding one common trend but Maximum Eigen value is founding two common trends. The same is the nature of conflict in cointegration of Bale Robe Nazareth and Mekelle at 18th lag. If the last two combinations are found to be under rule

of one price, it would be very questionable if their cointegration is strong enough as it is backed by one test and rejected by other. So it is logical step to focus next on appropriateness of the specification of the error correction model used, mainly the lag used to estimate the cointegration relationship.

Table 34 for model with restricted trend lag, normality and serial correlation testes for three markets

Markets	Lag used	Lag selection by different information criterions				Jarque - Bera test for Normality		LM - Serial correlation test	
		FPE	AIC	HQIC	SBIC	Statistic (6– df)	Prob	Statistic (135-df)	Prob
D–M–A	17	6	6	1	1	6.037	0.419	116.445	0.874
D–M–A	18	6	6	1	1	1.746	0.942	133.003	0.532
G–J–A	16	3	N/A	2	1	2.017	0.918	138.827	0.393
B–A–S	12	2	2	1	1	3.38	0.76	107.098	0.963
B–D–S	18	1	1	1	1	3.415	0.755	126.59	0.685
D–S–B	19	1	1	1	1	2.328	0.887	138.389	0.403
D–J–G	11	8	N/A	1	1	2.93	0.818	121.621	0.789
D–J–G	12	8	N/A	1	1	2.279	0.892	123.625	0.749
M–D–J	18	1	7	1	1	1.525	0.958	118.29	0.846
S–J–D	14	1	1	1	1	1.983	0.921	133.255	0.526
D–S–N	17	1	1	1	1	1.545	0.956	139.206	0.384
S–N–D	18	1	1	1	1	1.86	0.932	136.471	0.448
D–S–N	20	1	1	1	1	1.983	0.921	147.752	0.214
J–B–G	12	12	13	2	1	4.714	0.581	136.406	0.45
J–M–G	17	7	7	1	1	1.378	0.967	126.334	0.691
D-N-G	18	1	1	1	1	2.158	0.905	125.305	0.714
J–N–G	14	9	12	1	1	6.834	0.336	128.617	0.638
N–J–M	18	1	5	1	1	0.951	0.987	132.696	0.54
B–N–M	15	9	14	1	1	6.323	0.388	135.27	0.477

As can be seen from table 34 above, most of the market combinations at their fitted lag are found to be grossly over fitted. So the one common trend found on rank test is actually caused by weak power of the rank test when over fitted longer lags are used. Even for the single combination backed by some of the information criterions the lag used is very long; somehow implying that restricted trend is not appropriate for 3 dimension combination of markets. However assuming that there is trend on log of transaction cost data of those markets the best cointegration at 12 lag is found between deficit market of Jimma in South West, Surplus market of Bale Robe in South and another deficit market of Gonder at North. Their lags are backed by FPE but found to be under fitted by one lag based on AIC. Others are defining it as grossly over fitted with selection of 1 and 2 lags. Since there is no conclusive evidence about the inappropriateness of the lag fitted, it is logical to accept it as right fit. And for this combination the null of normality is not rejected at 58% level and the null of independence is not rejected at 44% level means their error vectors and terms are independently and normally distributed.

Two points can be understood from the combination of Jimma, Bale Robe and Gonder. First un-cointegrated market at lower dimension (Gonder) is integrating on higher dimensions to those markets which are cointegrated on lower dimension as expected on other two papers which followed relative routine. However the strongest cointegration observed between Addis Ababa and Mekelle is not extended in to third dimension, implying strength revisal. Strength reversal is observed from Mekelle and Addis Ababa to ward Jimma and Bale Robe which are adding Gonder. So this real data evidence that strength reversal is possible and the earlier methodologies are both theoretically and practically unsound at least up to third dimension. The second point is that distance did not seem to the determining factor for order of inclusion as Gonder is not the closest market either to Bale Robe or Jimma. And this goes against the conclusion of earlier papers.

However be relaxing the strict distributional assumptions made we can find markets which are cointegrated with one common trend. Table 35 and 36, below, shows that the combination of Bale Robe, Nazareth and Mekelle in one side, and the combination of Jimma, Gonder and Nazareth in other side are having lag lengths which are backed by AIC, but not others.

Table 35 for model with restricted trend lag, normality and serial correlation testes for three markets (with imperfect specification)

Markets	Lag used	Lag selection by different information criterions				Jarque - Bera test for Normality		LM - Serial correlation test	
		FPE	AIC	HQIC	SBIC	Statistic (6– df)	Prob	Statistic (135-df)	Prob
B-N-M	14	9	14	1	1	6.927	0.328	128.966	0.63
J-G-N	12	9	12	1	1	8.785	0.186	122.952	0.763

In both cases the null of normality at vector level is not rejected at 18% and the null of independence is not rejected at 36% level at each cumulative serial lag from 1 to 15. So even though they either have normality problem in some of their error terms or serial correlation problem at some serial lags, they are having normally and independently distributed error vector.

Table 36 for model with restricted trend the highest serial correlation for three markets (with imperfect specification)

Markets	LM statistics	Serial lags	Degree of freedom	Probability
B-N-M	7.3568	1	9	0.600024
J-G-N	38.2344	4	36	0.368307

Table 37, below, shows that for both market combinations the null of 0 and 1 rank are rejected with 99% confidence based on both maximum Eigen value and trace statistics. But the null of one common trend can't be rejected with 5% confidence. So in one side Bale Robe at south and Mekelle at north which are found to be weakly cointegrated at lower dimension are adding Nazareth at center. However Nazareth is not the closest market to Bale Robe or Mekelle and so distance did not seem to be the critical factor for

inclusion, again. Moreover new cointegration is developed between Deficit markets of Jimma at West and Gonder at North West with secondary central market of Nazareth at center.

Table 37 for model with restricted trend rank test for three markets (with imperfect specification)

Markets/ critical value	Lags	Trace statistics			Maximum Eigen value		
		0 rank	1 rank	2 rank	0 rank	1 rank	2 rank
B-N-M	14	79.409	33.237	5.4964	46.172	27.7406	5.4964
J-G-N	12	91.0002	38.4224	11.6466	52.5778	26.7759	11.6466
5%		42.44	25.32	12.25	25.54	18.96	12.52
1%		48.45	30.45	16.26	30.34	23.65	16.26

For the three combinations analyzed and found to have one random trend at higher lags time varying heteroskedasticity or ARCH effect is not found since the null of constant variance is not rejected at 12% level for all markets in each combination (see table 38, below).

Table 38 ARCH effect test for model with restricted trend at 3 dimensions

Combination	Lags	markets	Wald test		
			Wald statistics	Degree of Freedom	Probability
J–B–G	12	J	0.00	1	0.9585
		B	2.22	1	0.1358
		G	2.34	1	0.1265.
B–N–M	14	B	0.54	1	0.4614
		N	0.17	1	0.6773
		M	0.13	1	0.7232
J–G–N	12	J	0.02	1	0.8972
		G	0.31	1	0.5806
		N	0.00	1	0.9631

So in general given the fact that the selected combinations are using higher lags did imply the restricted trend model is not appropriate at third dimension. Indirectly it also implies that there is no trend in log of transaction cost of the country which can be caused by trended dynamics in macro economic variables like inflation, oil price, foreign exchange rate or/and macro policy in general. If there is trend in some areas like Mekelle or Dire Dawa it is related to trend in local conditions like transportation convince than macro economic parameters. However accepting the model as right fit, the strongest cointegration under rule of one price in third dimension is found between deficit market of Jimma at south west and Gonder at North West, with surplus market of Bale Robe at south. This is followed by two combinations which are revolving around secondary central market of Nazareth. One combination includes surplus market of Bale Robe and deficit market of Mekelle at North. Another combination is including two deficit markets of Jimma and Gonder. Given these facts let's extend the analysis in to fourth dimension next.

3.6. <u>Markets which are quadruple wise cointegrated under rule of one price</u>

3.6.1. <u>Restricted constant model</u>

If we keep strict normality and independence requirement in each error term and vector up to 10% level there is only one combination of 4 markets under rule of one price based on rank test; assuming average log of transaction cost is constant and the level data did not have any deterministic trend. Assuming that the selected lag of 16 is appropriate and error vectors are having white noise distributions there is one common trend in combination of deficit market of Jimma, Gonder and Mekelle and central market of Addis Ababa. This is so because the null of 0, 1 and 2 cointegration equations are rejected at 1% level but the null of single common trend can't even rejected at 5% level. Means the four markets are operating under rule of one price.

Table 39 for model with restricted constant rank test for four markets

Markets/ critical value	Lags	Trace statistics				Maximum Eigen value			
		0 rank	1 rank	2 rank	3 rank	0 rank	1 rank	2 rank	3 rank
M–J–A–G	16	165.24	76.46	38.26	8.17	88.78	38.2	30.09	8.17
5%		53.12	34.91	19.96	9.42	28.14	22	15.67	9.24
1%		60.16	41.07	24.6	12.97	33.24	26.81	20.2	12.97

And the error vectors are normally and independently distributed as there is no evidence of serial correlation or none normal distribution. Summarized serial correlation testes are given at cumulative serial lag of 15 and as can be seen from table 40 below the null of independently distributed error vector can't be rejected up to 82% level. Moreover each market's error vector is having Gaussian error term and the vector level J-B test shows that the null of normality can't be rejected up to 93% level.

Table 40 for model with restricted constant lag, normality and serial correlation testes for four markets

Markets	Lag used	Lag selection by different information criterions				Jarque - Bera test for Normality		LM - Serial correlation test	
		FPE	AIC	HQIC	SBIC	Statistic (8 – df)	Prob	Statistic (240 –df)	Prob
M–J–A–G	16	2	2	2	1	3.039	0.9319	219.32	0.83

Important question is related to the fact that if the observed lag order can be backed by any of the information criterions. To allow for better lag selection by information criterions given low degree of freedom the minimum of the maximum lag allowed is reduced from 10 to 5. And as can be seen above the modal lag selected for all maximum lags allowed from 5th to 18th lags is 1st or 2nd. Even up to 12th lag AIC is observed to select the 2nd lag but then after it start picking the maximum lag given and latter on around 18th maximum lag allowed all are observed to pick the maximum lag allowed. So the reduction of the maximum lag allowed from 10 to 5 did not cause the smallest lag

selected by all information criterions. And using the modal selections the combination is found to be grossly over fitted. So if we impose the strict normality condition it is not possible to find four markets which are cointegrated under rule of one price, under this model. What if we relax the strict normality and independence assumption?

Even when strict conditions are dropped it was not possible to find combination of four markets which are having white noise error vector. Many combinations were found to have one common trend with independently distributed error vectors and lags supported by information criterions. But most of them are having highly none normal distribution. So with the restricted constant model which allows for constant average log of transaction cost of creating space utility and which does not allow for trend or seasonal effect on level data, it is not possible to find four markets which are ruled by rule of one price. What about if these restrictions are dropped by using the model with unrestricted constant or restricted trend? These are dealt next.

3.6.2. **Unrestricted constant model**

Under restrict distributional assumptions and with out any restriction on their lags three market combinations are found to have one common trend. As can be seen from table 41 below the null of 0 to 3 cointegration equations are rejected with 99% confidence in both testes. However the null of one common trend can't be rejected at 5% level. So if the specification assumptions are attained these markets are ruled by rule of one price.

Table 41 for model with constant rank test for four markets

Markets/ critical value	Lags	Trace statistics				Maximum Eigen value			
		0 rank	1 rank	2 rank	3 rank	0 rank	1 rank	2 rank	3 rank
N-S-M-G	16	488.83	129.21	37.86	2.02	359.62	91.34	35.84	2.02
G-A-J-M	16	519.35	182.48	42.76	1.13	336.86	139.72	41.63	1.13
N-S-B-M	16	490.43	176.16	66.9	0	314.27	109.26	66.89	0
5%		47.21	29.68	15.41	3.76	27.07	20.97	14.07	3.76
1%		54.46	35.65	20.04	6.65	32.24	25.52	18.63	6.65

Moreover table 42 below clearly shows that there is no evidence of serial correlation or none normal distribution in their error vectors. The null of normality can't be rejected at 87% level for all combinations. And the null of independently distributed error vectors can't be rejected at 68% for all combinations. So assuming unrestricted constant model is appropriate, these market combinations at their used lags are having highly independently distributed error vectors. However the maximum lag selected by any of the information criterions, except once, is 3 and all are found to be grossly over fitted. Given the fact that rank test has weak power when inappropriately longer lags are used implies that there is no single common trend among this combinations. AIC result about combination of Gonder, Addis Ababa, Jimma and Mekelle is observed to select the maximum lag allowed at each level; so mode value was not able to be selected. This could imply the maximum lag is the better fit. But actually given the known tendency of AIC to select the

maximum lag at small samples and large dimensions with low degree of freedom, it is more probable it is caused by up ward bias of AIC.

Table 42 for model with unrestricted constant lag, normality and serial correlation testes for four markets

Markets	Lag used	Lag selection by different information criterions				Jarque - Bera test for Normality		LM - Serial correlation test	
		FPE	AIC	HQIC	SBIC	Statistic (8 – df)	Prob	Statistic (240-df)	Prob
N-S-M-G	16	1	2	1	1	3.796	0.875	224.508	0.756
G-A-J-M	16	3	N/A	1	1	1.92	0.983	200.622	0.97
N-S-B-M	16	3	3	1	1	2.537	0.96	229.068	0.683

So keeping the strict distributional assumptions intact there is no evidence of cointegration under one common trend in both restricted and unrestricted constant models. What if the restricted assumption are relaxed a little bit by demanding white noise error vector with possible none white noise error terms at equation and serial lags level?

Even though 10 different combinations are observed to have lags backed by some of the information criterions, 8 of them are observed to have series normality problem. But two combinations are observed to have normally distributed error vectors, with some evidence of serial correlation on their serial lags. The combination of Addis Ababa, Bale Robe and Jimma which was ruled by one price in third dimension is observed to add the deficit market of Gonder. At vector level and for selected order the null of normally distributed error vector can't be rejected at 12% level. Moreover the used lag order of 1 is also backed by both HQIC and SBIC. However at 9th serial lag highest cumulative serial correlation is observed, which is significant at 5% but not at 1% level. And at 15th lag the null of independently distributed error vectors is not rejected at 5% but 10% level (see table 43 and 44 below). To see the series ness of the serial correlation lags are selected by using information criterions on error terms and first lag of the four markets. And under the selected lag of 1 the null of independent error vectors is not rejected at 20% level. So even though there is some evidence of serial correlation the problem is not series enough to raise questions on the result of the rank test.

Table 43 for model with unrestricted constant lag, normality and serial correlation testes for four markets (with imperfect specification)

Markets	Lag used	Lag selection by different information criterions				Jarque - Bera test for Normality		LM - Serial correlation test	
		FPE	AIC	HQIC	SBIC	Statistic (8– df)	Prob	Statistic (240-df)	Prob
A-B-G-J	1	7	7	1	1	12.223	0.142	268.54	0.099
B-G-J-N	1	2	9	1	1	12.350	0.136	263.949	0.138

The implication is that the cointegration under rule of one price observed between Surplus market of Bale Robe, deficit markets of Jimma and Gonder, and central market of Addis Ababa did support the fact that other market at higher dimension are more

probable to be added to integrated market at lower dimension. This is so because the strongest cointegration is found between Addis Ababa, Jimma and Bale Robe at 3 dimensions. Unfortunately Gonder is not found in close distance to either market which rejects the assumption that distance is important factor for inclusion of markets in to system of one price. It seems inclusion have to do with distance but not with distance, per se.

Table 44 for model with unrestricted constant the highest serial correlation and serial lag selected by information criterions for four markets (with imperfect specification)

Markets	LM statistics	Serial lags	Degree of freedom	Probability
The lag with highest serial correlation				
A-B-G-J	176.9931	9	144	0.032056
B-G-J-N	151.6272	8	128	0.075558
The lag selected by information criterions				
A-B-G-J	20.2859	1	16	0.20762
B-G-J-N	19.0032	1	16	0.26850

Moreover the cointegration of Bale Robe, Gonder and Jimma under rule of one price is observed to add Nazareth, too. So the secondarily central market of Nazareth is also observed to be weakly cointegrated with surplus market of Bale Robe and deficit market of Jimma and Gonder. In this cointegration and at 8th serial lag there is some evidence of serial correlation at 10% level. But at 15th lag or 1st lag selected by information criterions there is no evidence of serial correlation. More over in all cases there is no evidence of none normality and the 1 lag used in VECM is backed by both HQIC and BSIC.

Table 45 for model with unrestricted constant rank test for four markets (with imperfect specification)

Markets/ critical value	Lags	Trace statistics				Maximum Eigen value			
		0 rank	1 rank	2 rank	3 rank	0 rank	1 rank	2 rank	3 rank
A-B-G-J	1	69.09	44.49	20.47	1.64	24.6	24.03	18.82	1.64
B-G-J-N	1	64.53	42.35	20.46	1.71	22.18	21.88	18.76	1.71
5%		47.21	29.68	15.41	3.76	27.07	20.97	14.07	3.76
1%		54.46	35.65	20.04	6.65	32.24	25.52	18.63	6.65

Table 45 above is giving rank test based on both trace statistics and maximum Eigen value. Based on trace statistics the null of 0 to 2 cointegration equations is rejected with 99% confidence but the null of one common trend is not rejected at 5% level. So based on trace statistics there is one common trend for both group of four markets. Means they are ruled by rule of one price. However the maximum Eigen value is giving contradicting results for both combinations. Based on maximum Eigen value the null of un-cointegration markets is not rejected at both 1% and 5% level, which means there is no cointegration between these markets. And this is impossible since at least we know that there is cointegration at lower dimension. And the null of 1 cointegration equation is rejected with 95% confidence but not 99% confidence. And the null of just two cointegration equation is rejected with 99% confidence for both combinations. But if there is 0 or 1 rank, it is not possible to have more than 2 ranks, which explains why the

result of the maximum Eigen value test is self contradicting. But the null of one common trend is not rejected at 5% level for both. So there is no strong evidence to contrary to doubt the existence of one common trend among these prices.

Table 46 ARCH test for 4 markets under unrestricted constant model

Combination	Lags	Markets	ARCH test		
			Wald statistics	Degree of freedom	Probability
A-B-G-J	1	A	0.09	1	0.7591
		B	0.06	1	0.8027
		G	1.11	1	0.2918
		J	0.76	1	0.3819
B-G-J-N	1	B	0.02	1	0.8810
		G	1.55	1	0.2127
		J	0.78	1	0.3756
		N	0.08	1	0.7834

ARCH effect is tested and for both combinations and it is found that the null of constant variance can't be rejected up to 21% level (see table 46, above). So given lack strong evidence to contrary it is logical to accept the conclusion that the two combinations of four markets are following one common trend or are governed by rule of one price.

Relatively the cointegration Bale Robe, Gonder, Jimma and Nazareth is much stronger than the cointegration of Addis Ababa, Bale Robe, Gonder and Jimma, if strength is measured by distribution of the error vectors. However at lower dimensions the combination of Addis Ababa, Bale Robe and/or Jimma was much stronger than any other combination. So strength reversal has not only theoretical but also practical possibility. And this will cast doubt on theoretical and practical strength of the methodology followed by Gonz´alez-Rivera and Helfand (2001) and Rashid (2006). Their methodology could identify the first combination but not the second stronger combination. But more robust conclusion can be only reached if higher dimensions are analyzed and the boundary of both combinations is fully mapped.

3.6.3. **Restricted trend Model**

To allow for possibility of trended log of transaction cost restricted trend model can be very useful. With in this model 7 different combinations are found to have one common trend and highly white noise error vectors. As can be seen from table 47 below, the null of 0, 1, and 2 ranks are rejected with 99% confidence by both testes. Moreover the null of one common trend can't be rejected at 5% level for all combinations and both testes. So assuming the specification assumptions needed for rank test are attained it is logical to conclude that these combination are ruled under one price. Means there is only common random trend which is keeping the prices together forming common rule of one price.

And this markets are selected given they are having strictly whites noise distribution and common trend. And table 48, below, clearly shows that the null of normality can't be rejected even at 29% level for all combinations. Moreover the null of independently

distributed error vector can't be rejected at 45% level for all combinations. So it is clear these combinations are having highly normally and independently distributed error vectors.

Table 47 for model with restricted trend rank test for four markets

Markets/ critical values	Lags	Trace statistics				Maximum Eigen value			
		0 rank	1 rank	2 rank	3 rank	0 rank	1 rank	2 rank	3 rank
A-J-M-G	13	142.76	82.24	37.82	8.92	60.53	44.42	28.9	8.92
G-A-D-M	13	175.78	86.38	35.54	6.59	89.4	50.84	28.96	6.59
A-D-G-M	14	239.93	112.33	42.77	3.8	127.6	69.56	38.97	3.8
S-A-N-J	12	130.74	70.02	38.42	9.5	60.71	31.61	28.92	9.5
M-D-A-B	14	207.46	79.67	33.79	8.3	127.79	45.88	25.49	8.3
B-M-S-N	15	417.16	150.61	34.07	5.62	266.55	116.54	28.45	5.62
S-J-D-N	14	146.62	68.84	35.89	11.87	77.78	32.94	24.03	11.87
5%		62.99	42.44	25.32	12.25	31.46	25.54	18.96	12.52
1%		70.05	48.45	30.45	16.26	36.65	30.34	23.65	16.26

The problem is that all combinations are over fitted based on all information criterions. And given the well known fact of lower power of rank test under such specification, it is very logical to conclude that there is more than one common trend in the cointegration of these markets.

Table 48 for model with restricted trend lag, normality and serial correlation testes for four markets

Markets	Lag used	Lag selection by different information criterions				Jarque - Bera test for Normality		LM - Serial correlation test	
		FPE	AIC	HQIC	SBIC	Statistic (8 – df)	Prob	Statistic (240-df)	Prob
A-J-M-G	13	7	7	1	1	1.916	0.983	211.971	0.904
G-A-D-M	13	6	6	1	1	2.535	0.96	192.342	0.989
A-D-G-M	14	6	6	1	1	6.131	0.633	217.024	0.854
S-A-N-J	12	1	1	1	1	4.108	0.847	231.177	0.647
M-D-A-B	14	1	5	1	1	4.563	0.803	239.671	0.494
B-M-S-N	15	3	3	1	1	2.966	0.936	242.052	0.451
S-J-D-N	14	1	1	1	1	9.557	0.298	218.464	0.837

So if strict normality and independence condition is demanded we can't find any market combination in four dimensions under rule of one price. And even when the strict normality and independence condition demanded on distributions of error vectors is relaxed; it is not possible to find market combinations which are sharing one common trend. So it is more logical to assume that the trend observed at some markets transaction is not related to national factors like inflation, foreign exchange rate, oil price and other macro factors but to local variables like the availability and quality of roads.

3.7. Five markets under rule of one price
3.7.1. Restricted constant model

Given the fact that estimation of lower dimension cointegration of markets which are cointegrated at higher dimension will have estimation bias, it is not logical to stop the search for markets under rule of one price at higher dimension because there is weak or no cointegration at lower dimension. This is so because the fact that estimation of cointegrated markets at higher dimension using combination of lower dimension combinations of these markets will introduce omitted variable bias (Gonz´alez-Rivera and Helfand 2001); and such biased estimation can possibly generate biased errors and can effect the estimated distribution of the error terms. As result not only lower dimension combination can be rejected as unfit based on serial correlation and normality test but also the rank test is not perfectly appropriate test for cointegration in case the errors vectors are not white noise.

Table 49 for model with restricted constant rank test for five markets

statistics	Rank	MARKET	critical values	
		J-A-N-D-S	5%	1%
Trace statistics	0	237.6714	76.07	84.45
	1	122.673	53.12	60.16
	2	66.8929	34.91	41.07
	3	28.7559	19.96	24.6
	4	0.8401	9.42	12.97
Maximum Eigen value	0	114.9984	34.4	39.79
	1	55.7801	28.14	33.24
	2	38.137	22	26.81
	3	27.9158	15.67	20.2
	4	0.8401	9.24	12.97

Keeping the strict distributional assumptions but with out any restriction on lag used on VEC model one combination of 5 markets is found to be under rule of one price. As can be seen from table 49 above the combination of deficit markets of Jimma in West and Dire Dawa in East are found to be to have one common trend with surplus market of Shashimiene at south central and central markets of Addis Ababa and Nazareth at center. The null of 0 to 3 ranks are rejected with 99% confidence by both testes. But the null of one common trend can't be rejected even at 5% level in both testes. So assuming the lag selected is appropriate there is one common trend between these five markets or they are ruled by rule of one price.

Table 50, below, also clearly shows that the null of independently distributed error vector can't be rejected at 34% level and the null of normally distributed error vector can't be rejected at 58% level. So the error vectors are normally and independently distributed. However when the lag fitted is compared with lag selected by information criterions it is

found to be grossly over fitted. As result there is no logical ground to accept the markets as following one common trend.

Table 50 for model with restricted constant lag, normality and serial correlation testes for five markets

Markets	Lag used	likelihood–ratio based lag selection				Jarque - Bera test for Normality		LM - Serial correlation test	
		FPE	AIC	HQIC	SBIC	Statistic (10 – df)	Prob	Statistic (375-df)	Prob
J-A-N-D-S	13	1	1	1	1	8.441	0.586	385.21	0.347

So keeping strict distributional assumptions and assuming un-trended level data with constant log of transaction cost there is no combination of 5 markets which are sharing one common trend. Even when strict normality for each error term and independence at each serial lag is relaxed it is not possible to find five combinations of markets which are sharing the same common trend with white noise error vector and lag supported by some of the information criterions. So estimation bias at lower dimension did not seem to be a critical reason for identification of cointegrated markets at lower dimension, at least up to fifth dimension under assumption of constant log of transaction cost and un-trended level data.

3.7.2. **Unrestricted constant model**

The combination of Jimma, Addis Ababa, Nazareth, Dire Dawa and Shashimiene are also found to be strongly cointegrated under rule of one price at 13th lag, even if we assume trended level data with constant log of transaction cost (see table 51, below). As was stated above there is no deterministic trend in level data to justify the model, but it will be useful to account for seasonal effect with out allowing for trended log of transaction cost. Table 51 below shows that based on maximum Eigen value and trace statistics the null of 0 to 3 cointegration equations are rejected with 99% confidence but the null of one common trend can't be rejected in either test at 5% level. So assuming the vector error correction model is correctly specified, it is logical to conclude they are operating under rule of one price. The basic question is do they have appropriate lag since this combination is selected given it is having highly white noise error vectors and terms.

This combination of markets at used lag is having highly white noise error vector, since at each serial lag the null of independence is not rejected at 10% level and at each error term the null of normality is not rejected at 10% level. For presentation convince, the summarized vector level results are given in table 52, below. And accordingly the vector level of null of normality is not rejected at 78% level and the vector level null of independence is not rejected at 70% level. However the model is found to be over fitted based on all information criterions, so it is logical to conclude that if we assume constant log of transaction cost there are no strongly cointegrated markets under rule of one price at 5th dimension. This is so if we assume trended or un-trended level data.

Table 51 for model with constant rank test for five markets

statistics	Rank	MARKET	critical values	
		J-A-N-D-S	5%	1%
Trace statistics	0	509.3211	68.52	76.07
	1	177.4704	47.21	54.46
	2	92.1505	29.68	35.65
	3	26.4116	15.41	20.04
	4	2.5702	3.76	6.65
Maximum Eigen value	0	331.8507	33.46	38.77
	1	85.3199	27.07	32.24
	2	65.7389	20.97	25.52
	3	23.8414	14.07	18.63
	4	2.5702	3.76	6.65

Table 52 for model with unrestricted constant lag, normality and serial correlation testes for five markets

Markets	Lag used	likelihood–ratio based lag selection				Jarque - Bera test for Normality		LM - Serial correlation test	
		FPE	AIC	HQIC	SBIC	Statistic (10 – df)	Prob	Statistic (375-df)	Prob
J-A-N-D-S	13	1	1	1	1	6.322	0.788	359.8493	0.704

However when the strict normality at each error term and strict independence at each serial lag demanded are relaxed to vector level demands, two combinations of five markets are observed to follow one common trend or to be governed by rule of one price. At second and third dimension the cointegration of Addis Ababa, Bale Robe and Jimma was found to be stronger than any other combinations; however in forth dimension the cointegration of Bale Robe, Jimma, Gonder and Nazareth was stronger than the cointegration of Addis Ababa, Gonder, Jimma and Bale Robe. So the strong cointegration at 4th dimension did not include Addis Ababa and this was taken as evidence of strength reversal. And the conclusion is also maintained in fifth dimension in which the first combination by adding Surplus market of Shashimiene was having better normality and independence in its error vectors (see table 53, below).

Table 53 for model with unrestricted constant lag, normality and serial correlation testes for five markets (with imperfect specification)

Markets	Lag used	likelihood–ratio based lag selection				Jarque - Bera test for Normality		LM - Serial correlation test	
		FPE	AIC	HQIC	SBIC	Statistic (10 – df)	Prob	Statistic (375–df)	Prob
B-J-N-S-G	1	2	6	1	1	13.012	0.223	370.503	0.556
G-B-J-A-S	1	2	6	1	1	15.308	0.121	379.025	0.432

The cointegration of surplus market of Bale Robe and Shashimiene with deficit market of Jimma and Gonder through secondarily central market of Nazareth has error vectors in which the assumption of normality is not rejected at 22% level and assumption of independence is not rejected at 55% level.

Table 54 for model with unrestricted constant the highest serial correlation for five markets (with imperfect specification)

markets	LM	Probability	lag	Degree of freedom
B-J-N-S-G	233.433	9	225	0.335779
G-B-J-A-S	247.9333	9	225	0.140613

But the relative figure for the second one which includes the three markets of Addis Ababa, Bale Robe and Jimma in addition to Gonder and Shashimiene is 12% and 14%, respectively. The point is that if we follow the recommendation of Gonz´alez-Rivera and Helfand (2001) we could identify the second combination of markets but not the first. This is so even though the cointegration is stronger in the first combination. For both combinations their lag is backed by HQIC and SBIC and there is no series serial correlation problem at any cumulative serial lag from 1st to 15th (see table 53 and 54 above). But the search method followed by Gonz´alez-Rivera and Helfand (2001) though found unsound in 5th dimension, it can only be rejected if both combinations are not have one common trend in higher dimensions. And there is evidence to show latter on that both are following one common trend.

One very important observation is that related to the fact that Shashimiene which the closest market to both Addis Ababa and Nazareth is included in 5th dimension in to the rule of one price but not before. This is clear indication that cointegration may be effected by distance but not by distance alone. So the search procedure can't follow road distance as recommended by Gonz´alez-Rivera and Helfand (2001) and applied by Rashid (2006).

Table 55 for model with unrestricted constant rank test for five markets (with imperfect specification)

statistics	Rank	MARKET		critical values	
		BJNSG	GBJAS	5%	1%
Trace statistics	0	108.339	105.2824	68.52	76.07
	1	66.0289	65.6843	47.21	54.46
	2	41.605	40.7827	29.68	35.65
	3	21.3111	20.0811	15.41	20.04
	4	1.6685	1.7859	3.76	6.65
Maximum Eigen value	0	42.3101	39.598	33.46	38.77
	1	24.424	24.9016	27.07	32.24
	2	20.2939	20.7016	20.97	25.52
	3	19.6425	18.2952	14.07	18.63
	4	1.6685	1.7859	3.76	6.65

As can be seen from table 55 above based on trace statistics for both combinations the null of only 0 to 3 cointegration equations are rejected with 99% confidence but the null of one unit root is not rejected even at 5% level. These means based on trace test the five markets are under rule of one price. And Based on maximum Eigen value the cointegration of Gonder, Bale Robe, Jimma, Addis Ababa and Shashimiene is also having one common trend at both 1% and 5% level. But the maximum Eigen value is less consistent in the cointegration of Bale Robe, Jimma, Nazareth, Shashimiene and Gonder. The null of only 0 and null of 3 cointegration equations is rejected with 99% confidence but not the null of only 1 and 2 cointegration equations at 1% and 5% level. So the test is contradicting it self since it is not possible to have 3 ranks with out having 1 or 2 ranks. So given lack of conclusive evidence to contrary the markets are accepted as following one common trend.

Before concluding the analysis ARCH effect is tested and for both combinations heteroskedasticity is observed only in two markets. As can be seen from table 56, below, for Shashimiene the null of constant variance is rejected at 1% level for both combinations. And for Gonder the null is rejected at 10% but not at 5% level. So there is evidence of heteroskedasticity mainly in Shashimiene and marginally in Gonder.

Table 56 ARCH test for 5 markets under unrestricted constant

Combination	Lags	Markets	ARCH test		
			Wald statistics	Degree of freedom	Probability
B-J-N-S-G	1	B	0.17	1	0.6820
		J	1.01	1	0.3153
		N	0.16	1	0.6896
		S	61.22	2	0.0000
		G	2.78	1	0.0954
G-B-J-A-S	1	G	2.76	1	0.0964
		B	0.12	1	0.7250
		J	1.06	1	0.3023
		A	0.04	1	0.8468
		S	39.14	2	0.0000

Focusing in first combination and Shashimiene the first coefficient is large and equal to 0.482993 and for Gonder it is 0.1988642 so there is logical reason to doubt if the cointegration of Shashimiene to Bale Robe, Jimma, Nazareth and Gonder under one common trend is resulted due to weak power of the rank test when there is strong ARCH effect. Note that the four markets excluding Shashimiene are found to be cointegrated under rule of one price with white noise error vector at fourth dimension. In the second combination the first coefficient of Shashimiene and Gonder are 0.4500022 and 0.2046188 respectively. And this are very large to cast doubt if the inclusion of Shashimiene in to cointegration of Bale Robe, Jimma, Addis Ababa and Gonder under rule of one price is also caused by the weak power of the rank test under ARCH effect. How ever ARCH effect in general is observed to have marginal effect in the Monte Carlo studies and since nulls are rejected at 1% level than 5% level, it is hardly possible to reject Shashimiene as un-cointegrated to the four markets.

3.7.3. <u>Restricted trend model</u>

If we accept 12th lag used for combination of Mekelle, Addis Ababa, Bale Robe, Nazareth and Shashimiene it seems they are strongly cointegrated under rule of one price with trended log of transaction cost. This is so because the null of 0 to 3 ranks are rejected with 99% confidence but the null of one common trend can't be rejected even at 5% level (see table 57, below). And this is the case based in both rank testes.

Table 57 for model with restricted trend rank test for five markets

statistics	Rank	MARKET	critical values	
		M-S-B-N-A	5%	1%
Trace statistics	0	292.8678	87.31	96.58
	1	138.1667	62.99	70.05
	2	84.5878	42.44	48.45
	3	40.2231	25.32	30.45
	4	5.2186	12.25	16.26
Maximum Eigen value	0	154.7011	37.52	42.36
	1	53.5789	31.46	36.65
	2	44.3646	25.54	30.34
	3	35.0045	18.96	23.65
	4	5.2186	12.52	16.26

However, even though they are having normally and independently distributed error vector (see table 58, below) the combination are found to be grossly over fitted at 12th lag based on all information criterions. And given weak power of rank test under such specification it is logical to conclude that there are more than 1 common trend between the customer center of Mekelle, surplus market of Shashimiene and Bale Robe, and central markets of Addis Ababa and Nazareth. So if the strong distributional assumptions are maintained, a single combination of markets can't be found with one common trend, which ever model is used.

Table 58 for model with restricted trend lag, normality and serial correlation testes for five markets

Markets	Lag used	likelihood–ratio based lag selection				Normality test		Serial correlation test	
		FPE	AIC	HQIC	SBIC	Statistic (10 – df)	Prob	Statistic (375-df)	Prob
M-S-B-N-A	12	1	1	1	1	1.973	0.996	370.9391	0.55

And even if the strict distributional assumptions are relaxed under restricted trend model, it was not possible to find five markets which are strongly cointegrated under rule of one price given they are having white noise error vector and lag used which is backed by information criterions. This will back the earlier conclusion reached that log of transaction cost at national level does not have either rising or declining trend, in period of 1996 to 2003.

3.8. <u>Six markets and above under rule of one price</u>

When log of transaction cost is assumed to be constant with out any control on seasonal effect and assuming un-trended level data under vector error correction model with restrict constant, it is not possible to find any market combination with one common trend; given the error vectors are white noise.

And even using vector error correction model with restricted trend which allows not only for possibility of trended log of transaction cost and deterministic trend in level data but also capacity to control seasonality by using centered seasonal dummy it is not possible to find six markets which are sharing one common trend at any possible lag. And for both restricted constant and restricted trend model six markets under rule of one price can't be found, even if the less restrictive distributional assumptions are maintained in which than error term level normality vector level normality and than independence at each serial lag, vector level independence at cumulative serial lag are demanded.

However when unrestricted constant is used assuming constant log of transaction cost and trended level data, which allows for controlling seasonal variation, it was possible to find three combinations of six markets cointegrated under rule of one price, with out restriction on lag used. These three combinations are discussed below.

3.8.1. <u>Unrestricted constant model</u>

Table 59, below, clearly show that based on both trace statistics and Maximum Eigen value the null of 0 to 4 cointegration equations are rejected with 99% confidence but the null of one common trend or the null that these combination of markets are governed by rule of one price can't be rejected at 5% level.

Table 59 for model with constant rank test for six markets

statistics	Rank	MARKET			critical values	
		G-S-N-B-J-D	S-G-A-N-J-M	G-M-S-A-D-J	5%	1%
Trace statistics	0	784.533	307.1224	536.5821	94.15	103.18
	1	406.6542	181.5819	265.1458	68.52	76.07
	2	226.6403	114.6545	168.9862	47.21	54.46
	3	102.4121	68.6518	87.2803	29.68	35.65
	4	25.3338	29.2184	25.9391	15.41	20.04
	5	1.2725	0.001	0.3885	3.76	6.65
Maximum Eigen value	0	377.8788	125.5405	271.4363	39.37	45.1
	1	180.0139	66.9274	96.1596	33.46	38.77
	2	124.2283	46.0027	81.7059	27.07	32.24
	3	77.0783	39.4334	61.3412	20.97	25.52
	4	24.0613	29.2174	25.5507	14.07	18.63
	5	1.2725	0.001	0.3885	3.76	6.65

So assuming that the lags used are appropriate it could imply that there is strong cointegration under rule of one price between these three combinations of six markets given in table 59, above. Given the fact that rank test is based on correct specification of the cointegration equation, the million dollar question is 'are the model used having an appropriate specification?'

In all three combinations the null of independently distributed error vectors can't be rejected up to 57% level (see table 60, below). Given the fact that these three combinations are selected to represent strongly cointegrated marketed by demanding that the null of independently distributed error vector not to be rejected at 10% level, on all serial lags from 1st to 15th , it is logical to conclude that there is no evidence of any serial correlation in all error vectors. The null of normality is also not rejected even at 64% for all combinations and given for each error term the null of normality is not rejected at 10% level the implication is that the error vectors are strongly normally and independently distributed as demanded by rank test.

Table 60 for model with restricted trend lag, normality and serial correlation testes for six markets

Markets	Lag used	likelihood–ratio based lag selection				Normality test		Serial correlation test	
		FPE	AIC	HQIC	SBIC	Statistic (12 – df)	Prob	Statistic (540-df)	Prob
G-S-N-B-J-D	11	1	1	1	1	4.2	0.979	532. 9	0.578
S-G-A-N-J-M	10	1	6	1	1	9.6	0.65	522.2	0.701
G-M-S-A-D-J	11	1	1	1	1	3.9	0.985	531.5	0.595

However the lag fitted is found to be much larger than the lag selected by all information criterions. Under such over fitted cointegration vectors Monte Carlo evidence did show that rank test have tendency to find cointegration when there is no real cointegration. So it is logical to conclude that there are no six markets which are following one strong common trend or strongly governed by rule of one price. And even if relaxed assumptions are demanded on the distribution of the error vectors it is not possible to find six markets under rule of one price.

3.8.2. <u>More than six markets under rule of one price</u>

Even if 7 and 8 markets are considered it was not possible to find markets under rule of one price for both restricted distributional requirement and relaxed distributional requirements and all models. So in general which ever model is used it is not possible to find six or more markets which are sharing one common trend or which are governed by rule of one price. The logical conclusion that could follow will be that the whole country is not ruled by one price but there are two combinations of five markets which are sharing two different single common trends. How ever analysis in common trend below will show that both combinations are sharing the same common trend. The reason the combinations are not cointegrating in to one common trend based on rank test is because their error vectors was found to be highly none white noise at first lag. This is clear evidence that rank test which is developed for varibles with white noise errors may be is inadequate in all cases. More over the search procedure followed by Gonz´alez-Rivera

and Helfand (2001) will not fail to find the strongly cointegrated markets under rule of one price, if information on common trend given below is considered. But if analysis for Ethiopian wheat markets was based only on rank test, only, it will end up in inconclusive result. Even though their methodology is convenient for application, it is not theoretically sound. Fortunately for Ethiopian data used there would be no difference on conclusion, if their methodology or the more tedious search for all possible permutations, as done in this paper is used. How ever distributional assumptions have to be tested at each stage even in their methodology.

Unfortunately the use of distance in order of inclusion recommended by Gonz´alez-Rivera and Helfand (2001) and used by Rashid (2006) is not sound in under developed market economy like Ethiopia with high market failures and dysfunctional institutions. And this is clearly seen in the above analysis. Coming to Ethiopian wheat market before possible recommendations are given, let's try to grasp the full picture by estimating the common trend and the short run dynamics of the markets.

4. <u>Dynamic analysis of short run parameters</u>

Even though cointegration analysis and the identification of the single common trend and its determinants will give us information about equilibrium level relationship between the market prices in the long run, it may be poor guide for short run policies. In practical world not only the long run equilibrium but also the short run dynamics of a system needs to be understood. In this paper short run dynamics is analyzed by considering persistence of system wide shock and statistical significance of adjustment parameters. The persistence analysis is based on Pesaran and Shin (1996) methodology which uses normalized change in forecast variance. This will measure the percentage of the initial shock that is persisting in each month (period). Persistence is function of all parameters in vector error correction model but it will be also informative to analyze the sign and statistical significance of the adjustment parameters. If an adjustment parameter of the first market related to the first cointegration relation is -0.5 and significant, it means 50% of the deviation from equilibrium in the first cointegration relation will be corrected in first period by the first market. Or in other words the market will take two periods (1/0.5 = 2) to correct for disequilibrium initiated in the first cointegration relation. Moreover if a market did not react to any deviation from equilibrium in all cointegration equations it is said to be weakly exogenous. And logically markets which are weakly exogenous are the preferable place for intervention since what ever shock created there is corrected by other markets.

For analysis of the short run dynamics of Ethiopian wheat markets the final 2 combinations of 5 markets are considered. According to Pesaran and Shin (1996) the persistence profile can behave in short run as integrated or I(1) series but in long run (as $t \rightarrow \infty$) it will decline to zero. And persistence analysis will map the dynamics of the shock until it decays to ward zero. Persistence profile is insensitive to market order for given identification assumption and this is the advantage of the profile over impulse response function. However it is sensitive to the identification assumption imposed in the cointegration vectors. So for better information two groups of persistence profile are

fitted i.e. one for identification assumption based on central market and one for identification assumption with highest probability of normality. The first one is more informative but the second one is more appropriate so both are used in analysis below.

4.1. <u>Short run dynamics between Addis Ababa and other 4 markets ruled by one price</u>

Figure 8 below clearly shows that for all market cointegration the impact of system wide shock will disappear in less than 10 months. But the deficit market of Gonder seems to have very weak cointegration with central market of Addis Ababa as it is observed to have higher persistence than any other markets. The combination of Deficit market of Jimma in west and central market of Addis Ababa is having the strongest cointegration with quickly disappearing impact of system wide shock compared to any other combination in the system of one price. This may imply that strongest cointegration in higher dimension are first to be identified by the rank, normality and serial correlation test in lower dimension. This is so because Gonder is included in to the system of one price in 4th dimension but Jimma and Bale Robe are introduced in to the system in second dimension and the latter are having low persistence or stronger cointegration in 5th dimension. If so this will back the search procedure used by Gonz´alez-Rivera and Helfand (2001) and Rashid (2006). Unfortunately such conclusion did not seem to be general as Shashimiene which is found to be cointegrated to this rule of one price only in 5th dimension is having strong cointegration compared to both Bale Robe and Gonder. So it is possible strongly cointegrated market in higher dimension when analyzed in lower dimension with its biased coefficients may not show cointegration. Means strongly cointegrated markets in higher dimension are not necessarily the one which show cointegration with white noise error vector at lowest possible dimension.

Figure 8 Persistence between Addis Ababa and other 4 markets ruled by one price 1
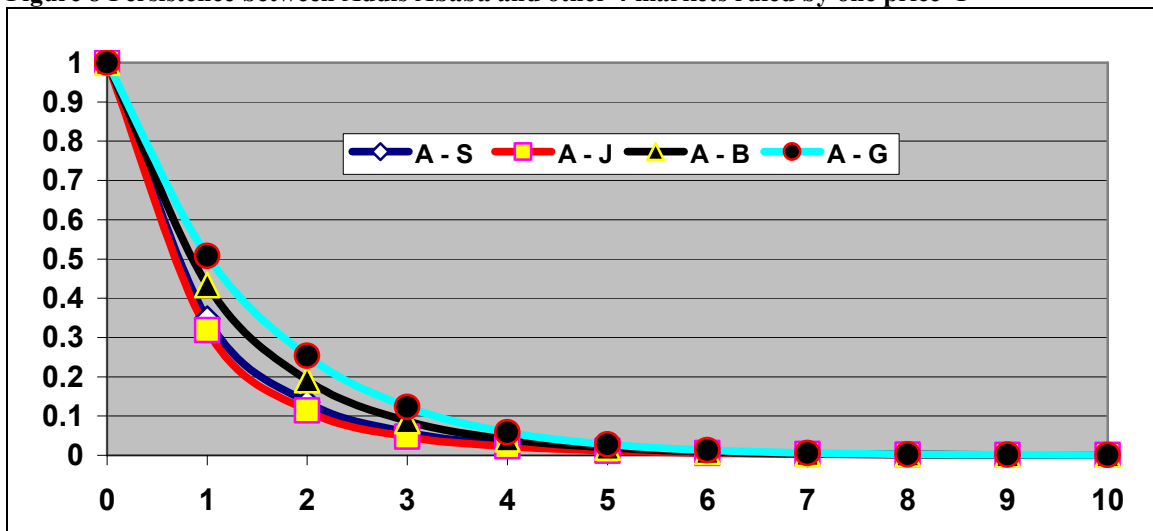
To summarize the persistence over 30 months two statistics are used. The first one is mean persistence which is the weighted average of the months it takes for the shock to eliminate it self; the weights being the shocks persisting in each month. The second

statistics is the persistence observed in first month. Gonzalo and Pitarakis (2000) used median persistence and mode persistence. Median persistence is the median time from 0 to the elimination of the impact of the shock and mode persistence is the time period where 50% of the shock is eliminated. But for the first one there is need to test if the persistence is statically equal to zero at each period. Pesaran and Shin (2006) did show that persistence profile has Gaussian (normal) distribution, so we can use normal tables to test for statistical significance. However the calculation of persistence variance is tedious and there is no much additional information that can be gained by analyzing median persistence to warrant the additional effort. For mode persistence its use is avoided because for most market combinations more than 50% of the shock is eliminated in just one month. So it is preferable for this study, if monthly persistence is used to measure the level of persistence in first month following the shock.

Table 61 Summery statistics for Persistence between Addis Ababa and other 4 markets ruled by one price 1

Market combination	A – S	A – J	A – B	A – G
Mean persistence in months	0.663446	0.606972	0.815844	0.964209
Mean persistence in days	19.90338	18.20916	24.47532	28.92627
First month Persistence (%)	0.34834	0.318259	0.433251	0.507999

And as can be seen from table 61 above on average it will take 0.6 of a month or 18 days for deviation from equilibrium between deficit market of Jimma and central market Addis Ababa to be corrected. And in first month 68% of the shock is eliminated in the system. However In case of deficit market of Gonder at North West and central market of Addis Ababa only less than 50% of the shock is eliminated in first month. And on average it will take 0.96 of a month or 29 days for the shock in the system to eliminate it self. The cointegration of Addis Ababa and Surplus market of Bale Robe at South will take around 24 days or 0.81 of a month to eliminate system wide shock in their long run relationship. And in first month around 43% of the shock will be still persisting. The second strong cointegration is observed between Surplus market of Shashimiene at south central and Addis Ababa. It will take close to 20 days on average for shock to eliminate it self and on first month 65% of the shock will be eliminated. So in terms of strength of cointegration Addis Ababa and Jimma seems to be to have the strongest cointegration followed by Addis Ababa and Shashimiene. And the weakest link in the system of one price is between Addis Ababa and Gonder followed between Addis Ababa and Bale Robe. Both Jimma and Shashimiene are located in coffee exporting region and zone of the country. Moreover Shashimiene is also a central hub for Chat trade in southern part of the country. Given coffee and Chat are the two most important export commodity of the country and given the experience, network, information and capital gained in coffee and Chat market can be used in grain trade, it may imply that the better capital, information, business network and experience gained by coffee and chat traders may be resulting in better cointegration of these markets than distance per se. When the markets are functioning well distance may be the main factor for determining the boundary for the rule of one price but when there are huge market failures other factors behind distance which are related to the level of market failure and the availability of complementary institutions to solve the market failure problem may determine the number of market under rule of one price.

Shocks between two markets which are pair wise cointegrated under larger dimension are corrected in the system not by the adjustment in price of the two markets, alone. So each market can potentially correct shocks initiated in any other market, if both are operating under the common rule of one price. This can be analyzed be observing the adjustment parameters of the vector error correction model given in table 62 below. Given the observed ARCH effect on Gonder and Shashimiene robust standard errors are used to improve the efficiency of the estimators.

As can be seen in table 62 below the deviation from equilibrium in long run relationship between Addis Ababa and Shashimiene is not adjusted in Shashimiene, which is observed to be weakly exogenous to any shock emanating from all cointegration relations. However when shocks are emanating form Shashimiene all markets but Gonder are reacting in the wrong direction to amplify the shock than to eliminate it as expected in cointegrated network of markets. Moreover when there is shock between Addis Ababa and Bale Robe it is again observed to be amplified in Addis Ababa. So the central market of Addis Ababa is amplifying if shocks are initiated in surplus markets and when the surplus market is Shashimiene most markets are going to confusion and disarray to amplify the shock.

Table 62 Adjustment parameters in rule of one price which include Addis Ababa and other 4 markets 1

	Cointegrated market with Addis Ababa under 5th dimension				LM test for over all significance (df-4)	
	Shashimiene	Jimma	Bale Robe	Gonder	Statistics	Prob.
Shashimiene	-0.091	-0.016	0.055	0	1.25	0.8697
Jimma	0.4***	-0.468***	0.033	-0.142	27.69	0
Bale Robe	0.35*	-0.065	-0.197*	0.107	10.89	0.0278
Gonder	0.161	0.018	0.004	-0.333***	18.57	0.0010
Addis Ababa	0.284**	-0.043	0.131*	-0.048	21.77	0.0002

Note 5*, ** and * imply significance at 1%, 5% and 10% level, respectively.**

Again the strongest cointegration between Jimma and Addis Ababa is also backed by adjustment parameters which is showing that it will take in Jimma close to two months or 2.14 month to be precise to correct shock initiated in long run cointegration relation between Jimma and Addis Ababa. The relative figure for Bale Robe is 5 months and is not significant at 5% but 10% level. For Gonder it will take 3 months to correct shocks initiated in the long run relationship between Addis Ababa and Gonder. So in general the customer centers of Jimma and Gonder have better information flow with Addis Ababa than surplus markets of Shashimiene and Bale Robe. The market seems to face confusion when shocks are initiated in surplus markets than deficit markets.

However since the identification assumption and normalization selected can have impact on the estimated parameters based on canonical correlation the identification assumption and normalization with highest probability of normality is found, if the markets are normalized against Shashimiene. And the same short run dynamics are analyzed for identification assumption against Shashimiene and are given in table 63 below.

Table 63 Adjustment parameters in rule of one price which include Addis Ababa and other 4 markets 2

	Cointegrated market with Shashimiene under 5th dimension				LM test for over all significance (df-4)	
	Gonder	Bale Robe	Jimma	Addis Ababa	Statistics	Prob.
Gonder	-0.334***	0.004	0.018	-0.005	18.57	0.0010
Bale Robe	0.107	-0.197*	-0.065	-0.163	10.89	0.0278
Jimma	-0.142	0.033	-0.468***	0.041	27.69	0.0000
Addis Ababa	-0.048	0.131*	-0.043	-0.437***	21.77	0.0002
Shashimiene	0	0.055	-0.016	0.053	1.25	0.8697

Note 6*, ** and * imply significance at 1%, 5% and 10% level, respectively.**

When ever there is shock between Shashimiene and any other market combination it is corrected on the other cointegrating market not in Shashimiene, which is found to be weakly exogenous. Gonder is correcting shock in its long run relationship with Shashimiene in 3 months. But fast correction is observed by Jimma and Addis Ababa which are correcting shock observed in their long run relationship with Shashimiene in close to 2 months or 64 days and 69 days, respectively to be precise. The slowest correction is observed in long run relationship between Bale Robe and Shashimiene which is observed to take around 5 months. Moreover in this relationship their shock is observed to be amplified by reaction of Addis Ababa in wrong direction. So it seems the market can better process and adjust shocks emanating in customer center of Jimma and Gonder and central market of Addis Ababa than producer centers of Bale Robe or Shashimiene. And it is logical given non tradable nature of the wheat and more or less marginal increase in per capital income of the population demand will be very predictable than supply which is dependent on random natural shocks. And given under developed market as observed by traders' lack of access to capital, proper storage and accurate information needed for effective temporal and spatial arbitration (Eleni 2001 and Eleni et al 2003), random shocks initiated in less predictable supply side than shocks initiated in more stable demand side may not be easy to correct. Actually the central market of Addis Ababa is observed to amplify the shocks observed in producer markets of Bale Robe and Shashimiene. And given the deficit market of Mekelle and Dire Dawa which are having highly aid dependent population are not in the rule of one price will also avoid the random shocks that can be initiated from distribution of wheat for free in form of aid. Both Gonder and Jimma are food deficit but they are neither drought prone areas nor big towns with large slumps. So food aid is less important in these markets than in Mekelle or Dire Dawa.

And as can be seen from persistence profile in figure 9 and table 64 below there is strongest cointegration between customer center of Jimma and Surplus market of Shashimiene. In which in just first month 74% of system wide shock will be corrected and on average it will take 12 days to correct it. This is higher than the correction observed by Jimma it self which is correcting 47% of its own shock in one month. So the over all marketing chain is also contributing to farther stabilization of this market. The next strong cointegration is observed between central market of Addis Ababa and Surplus market of Shashimiene; in which system wide shocks are corrected on average of 17 days in which 68% of the shock is corrected in just one month. Again this is higher than the

own adjustment of Addis Ababa, which is observed to correct around 44% of the shock in one month, if it is initiated in its own cointegrating relation with Shashimiene.

Figure 9 Persistence between Addis Ababa and other 4 markets ruled by one price 2
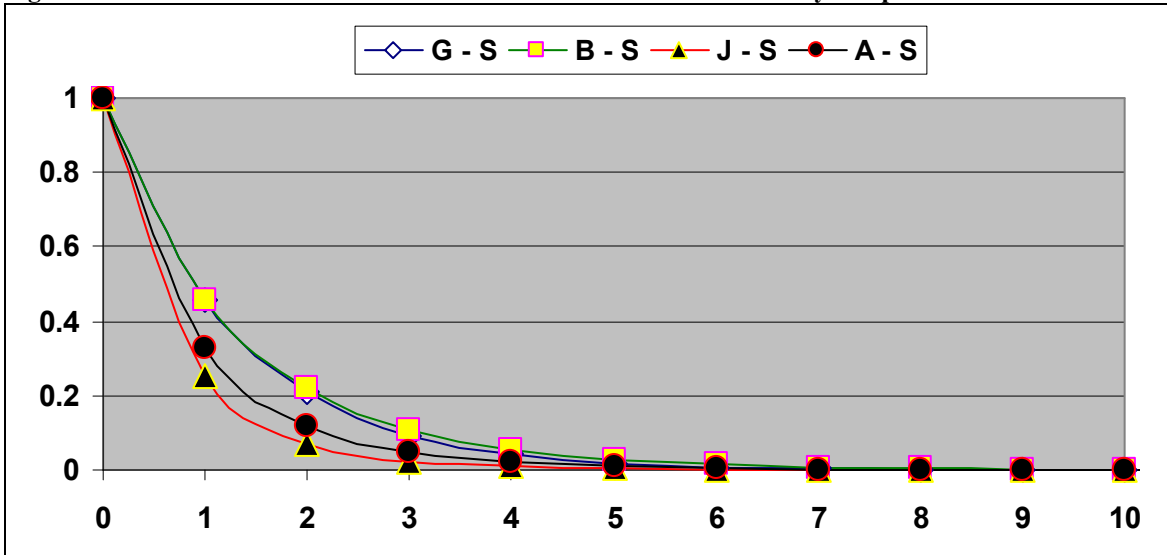
The slowest adjustment is observed between Bale Robe and Shashimiene, in which 45% of the system wide shock will be corrected in first month and on average it will take 29 days to correct a system wide shock. So even though Addis Ababa is amplifying the shock of these markets and they are very slow in correcting their own shock the over all net work of markets which are found under rule of one price is quickly sorting and correcting any system wide shock.

Table 64 Persistence between Addis Ababa and other 4 markets ruled by one price 2

Market combination	G – S	B – S	J – S	A – S
Mean persistence in months	0.820706	0.95324	0.409077	0.576494
Mean persistence in days	24.62117	28.5972	12.27231	17.29483
First month Persistence (%)	0.455524	0.456875	0.253098	0.328239

The same is the case in case of Gonder which is observed to have first month correction of 45% but its own correction is around 33% and in general it will take around 24 days on average to correct any shock observed in its long run relationship with Shashimiene. So for both order of variables and choice of normalization it is clear that the strongest cointegration is between Jimma, Addis Ababa and Shashimiene. The reason Shashimiene was not found to be strongly cointegrated with other markets in lower dimensions is because it is weakly exogenous to those markets. And the shocks initiated in Shashimiene are corrected by the system of other markets and until this system is fully built in higher dimension there was specification bias. And given biased estimation in lower dimension it was not possible to find cointegration with white noise error vectors between Shashimiene and other markets in lower dimension. Again these is another evidence about the fact that it is not necessarily true that strongly cointegrated markets at higher dimension will be found to have strong cointegration with white noise error vectors in lower dimension.

4.2. <u>Short run dynamics between Nazareth and other 4 markets ruled by one price</u>

The cointegration of Nazareth with the four markets of Jimma and Gonder as deficit markets and Shashimiene and Bale Robe as surplus markets is much stronger than the cointegration of Addis Ababa and the above four markets; if strength is measured by white noisiness of the error vector. The market combination which includes Nazareth in to the four markets of Jimma, Gonder, Shashimiene and Bale Robe; but which excludes Addis Ababa is dealt below. For presentation convince Nazareth is chosen as normalizing variable given its importance in both export and import market.

When Nazareth is used as central market the strongest cointegration is observed first between Nazareth and Jimma followed by Nazareth and Shashimiene (see figure 10 below). Moreover Gonder and Bale Robe again are having the weakest cointegration in the rule of one price.

Figure 10 Persistence between Nazareth and other 4 markets ruled by one price 1
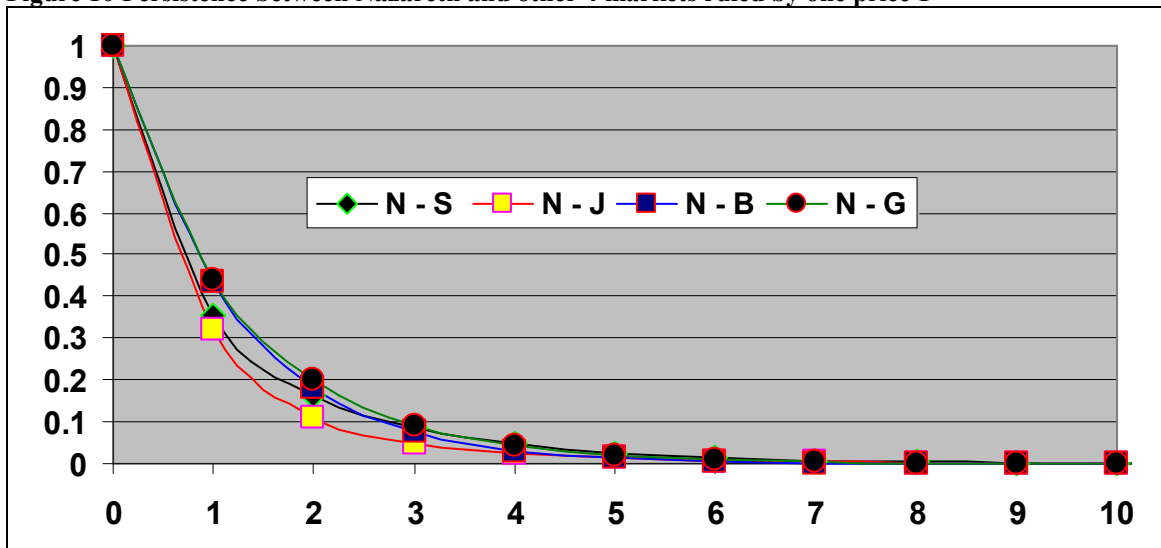
In both Jimma and Shashimiene integration with Nazareth when a system wide shock is injected to these markets, it is corrected on average of 18 and 25 days, respectively. But the relative figure for Bale Robe is 22 days and for Gonder is 25 days. One interesting point about Shashimiene is that even though 65% of the system wide is shock is corrected in first month on average it is observed to take 25 days to eliminate its impact. Means the remaining shocks are having long memory and will take time before their impact can be eliminated. This is consistent with ARCH effect observed in Shashimiene. Which also implies that since shocks are clustered stabilization of price if targeted in Shashimiene will be very costly and complicated.

Table 65 Persistence between Nazareth and other 4 markets ruled by one price 1

Market combination	N – S	N – J	N – B	N – G
Mean persistence in months	0.835542	0.615152	0.730741	0.831335
Mean persistence in days	25.06626	18.45456	21.92223	24.94005
First month Persistence (%)	0.353012	0.320257	0.437202	0.440079

For Jimma the first month correction is 67% but for both Bale Robe and Gonder the relative figure is 56%. So in one month both Jimma and Shashimiene are correcting significant slice of a system wide shock. But then after the correction process is very sluggish in Shashimiene. Bale Robe and Gonder are having the higher persistence compared to Jimma at all lags and to Shashimiene at first month. But as was stated before these shocks are corrected in the system of one price possibly by all market; so which market is correcting which shock is an important area which can be observed from the adjustment parameters. In the next analysis robust standard errors are used to account for ARCH effect in the estimation process.

And table 66 below shows that Shashimiene is weakly exogenous even in this group. This is so since it is not correcting any deviations from equilibrium resulting in all cointegration equations. Moreover Jimma and to some extent Bale Robe are observed to amplify any disequilibrium resulting on the cointegration relation between secondary market of Nazareth and Shashimiene. Jimma and Gonder did correct any deviation from equilibrium in their long run equilibrium with Nazareth in 2 and 3 months, respectively. As before strong correction of shocks is observed in deficit market of Jimma followed by another deficit market of Gonder. Both surplus centers of Shashimiene and Bale Robe are not correcting deviations or shocks on their long run relationship with Nazareth.

Table 66 Adjustment parameters in rule of one price which include Nazareth and other 4 markets 1

	Cointegrated market with Shashimiene under 5th dimension				LM test for over all significance (df-4)	
	Shashimiene	Jimma	Bale Robe	Gonder	Statistics	Prob.
Shashimiene	-0.189	-0.025	-0.009	0.026	4.63	0.3274
Jimma	0.431***	-0.458***	0.058	-0.151	27.77	0
Bale Robe	0.348*	-0.089	-0.198	0.105	10.12	0.0384
Gonder	0.106	0.008	-0.033	-0.319***	19.85	0.0005
Nazareth	0.222	-0.042	0.149	-0.039	10.67	0.0305

Note 7*, ** and * imply significance at 1%, 5% and 10% level, respectively.**

In terms of weak exogenous-ness only Shashimiene is clearly weakly exogenous since the null of weak exogenous can't be rejected at 32% level. For others the null is rejected at 5% level and mainly for Jimma and Gonder it is rejected at 1% level. But to have more accurate result and to identify the impact of identifying assumption on the above conclusion, lets use market order or identification assumption with highest probability of normality from all possible identification assumptions.

When the identification assumption with highest normality is selected the major difference observed is related to coefficients of Nazareth. Now Nazareth is observed to amplify shocks initiated in cointegration relationship between Bale Robe and Gonder. Moreover Bale Robe and Jimma are also observed to amplify shocks initiated in cointegration relation between Gonder and Shashimiene. However the system is kept in tact by corrections made by Jimma and Nazareth. Jimma will correct in nearly 2 months or 66 days any deviations in its equilibrium relationship with Gonder and the relative figure for Nazareth is 78 days. In terms of weak exiguousness still there is no difference in conclusion in which only Shashimiene is clearly weakly exogenous.

Table 67 Adjustment parameters in rule of one price which include Nazareth and other 4 markets 2

| | Cointegrated market with Shashimiene under 5th dimension | | | | LM test for over all significance (df-4) | |
	Bale Robe	Jimma	Nazareth	Shashimiene	Statistics	Prob.
Bale Robe	-0.198	-0.089	-0.126	0.348*	10.12	0.0384
Jimma	0.058	-0.458***	-0.036	0.431***	27.77	0
Nazareth	0.149*	-0.042	-0.384***	0.222	10.67	0.0305
Shashimiene	-0.009	-0.025	0.239	-0.189	4.63	0.3274
Gonder	-0.033	0.008	0.107	0.106	19.85	0.0005

Note 8*, ** and * imply significance at 1%, 5% and 10% level, respectively.**

However the choice of normalization is observed to make difference in persistence of system wide shocks. Any system wide shock will take much longer time to eliminate it self in cointegration relation between Jimma and Gonder than any other market combination. So Jimma even though is quickly adjusting to its own shocks initiated in its long run relationship with Gonder it will be adversely effected by shocks initiated in other markets and as result it will need longer time to sort them out. This is in line with observations in adjustment parameters in which other markets except Jimma and Nazareth are none response to shocks or reacting in wrong direction. However things are sorted by adjustment made in Nazareth and Jimma.

Table Figure 11 Persistence between Nazareth and other 4 markets ruled by one price 2
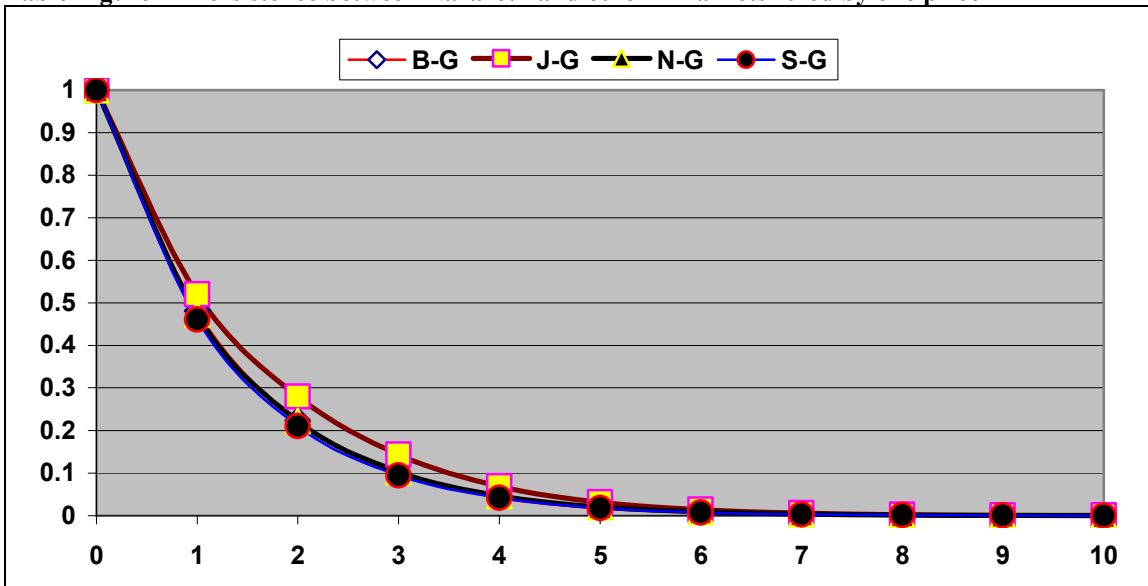
Other markets are having persistence which is hard to rank by visual inspection means they are having more or less the same persistence. So let's focus on summery statistics given in table 68 below.

And as can be seen from table 68 below the cointegration between Gonder and three markets of Bale Robe, Nazareth and Shashimiene are having similar persistence with persistence at first month equal to 48%, 47% and 46% of the initial shock, respectively. And for three of them the mean persistence is hovering around 25 to 26 days. However in

case of the cointegration of Gonder and Jimma only 48% of the initial system wide shock will be eliminated in first month and it will take more than one month on average to eliminate system wide shock.

68 Persistence between Nazareth and other 4 markets ruled by one price 2

Market combination	G– B	G – J	G – N	G – S
Mean persistence in months	0.833749	1.018091	0.859286	0.820989
Mean persistence in days	25.01246	30.54272	25.77858	24.62967
First month Persistence (%)	0.480866	0.520449	0.472248	0.461104

4.3. <u>Conclusion on short run dynamics</u>

So the country has two common trends which include five markets each. In which two are surplus and two are deficit markets mediated by two central markets of Nazareth in one direction and Addis Ababa in other direction. And to make comparison between them 6 different statistics will be used. Four are related to the minimum and maximum observed persistence in the combination of markets measured by first month persistence and mean persistence in days. Another two are related to the average of the persistence in the four cointegration relationships of 5 markets based on the two persistence measures stated above.

Table 69 the level relative strength of cointegration between the two market combinations

Identification With	Measures	Combination which includes Addis Ababa			Combination which includes Nazareth		
		Mean	Min	max	Mean	Min	max
Central market	Mean days	22.879	18.209	28.926	22.596	18.455	25.066
	1st persistence	0.402	0.318	0.508	0.388	0.32	0.44
Highest normality	Mean days	20.696	12.272	28.597	26.491	24.63	30.543
	1st persistence	0.373	0.253	0.457	0.484	0.461	0.52

Table 69 above shows that comparison is dependent on selection of identification assumptions imposed or market order used. If the analysis is done either against central market of Addis Ababa or Secondarily central market of Nazareth it seems the combination that includes Nazareth is having better cointegration. Both combinations are having more or less the same minimum persistence with 18 days in terms of average days and 32% in terms of first month persistence. However the combination which includes Addis Ababa is having higher maximum persistence with 29 days in terms of average days and with 51% of first month persistence compared to 25 days and 44% of Nazareth, respectively. Means Gonder and Bale Robe are adjusting quickly in market combination which includes Nazareth than in combination which includes Addis Ababa. As result the mean persistence is observed to be lower in the combination which includes Nazareth. In which in the combination which includes Addis Ababa the average first month persistence is 40% but in combination which includes Nazareth it is 39%. So even though the difference is not big the combination which includes Nazareth is marginally more

cointegrated than the combination which includes Addis Ababa. The same is observed if average days than first month persistence are used.

However if identification assumption with highest probability of normality is selected the combination which includes Addis Ababa will have the lowest persistence which ever measure is used. This is related to better cointegration between Addis Ababa and Jimma in the first combination and long persistence between Jimma and Gonder in the second combination. Still the combination which includes Addis Ababa is having lower mean and minimum persistence based on both measures, if compared with both identification assumptions on combination that include Nazareth.

The logical conclusion has to be identification assumption dependent and out of 20 permutations of markets or identification assumptions 1 is relevant and 19 are redundant and it is logical to select the identification assumption or market order which can result on higher normality. So even though the error vectors which include Nazareth have more normal error vector in their vector error correction model, the one which includes Addis Ababa is more cointegrated at margin than the one which includes Nazareth.

5. <u>Identification of single common trend and its determinants</u>

Up to this point markets which are ruled by one price are identified and their short term dynamics is analyzed. The final step for robust policy recommendation needs the estimation of the single common trend which is keeping these markets together. Moreover identification of markets which are having major impact on this common trend will result on identification of markets where effective policy intervention can be done with least possible transaction cost. The methodology used is based on permanent and transitory decomposition of cointegrated vectors developed by Gonzalo and Granger (1995).

In the analysis below F_A means the common trend which is cointegrating the five markets which include the central market of Addis Ababa. F_N means the common trend which is keeping the five markets which include Nazareth under rule of one price.

$$F_A = -4.3534381A - 9.2153955B - 6.1413652J - 4.8007328G + 11.096666S$$

As can be seen above in the rule of one price which includes Addis Ababa the first and second main determinants of the common trend are producer centers of Shashimiene and Bale Robe, respectively. The next strong impact is coming from deficit market of Jimma. Both Gonder and Addis Ababa are having the lowest impact on common trend, respectively.

Table 70 below clearly shows that the impact of both Addis Ababa and Gonder in the common cointegrating trend is statistically insignificant. The economic implication is that the deficit market of Gonder and the central market of Addis Ababa are price takers as the common trend is not determined by them. Means the market clearing price is mainly

discovered in surplus markets of Shashimiene and Bale Robe and one deficit market Jimma in the period of 1996 to 2003.

Table 70 the statistical significance of impact of markets in the rule of one price or common trend on market combination which includes Addis Ababa

Null	L_P	DF	probability
$a_{\perp A}=0$	1.398923658	1	0.236904031
$a_{\perp G}=0$	1.385296	1	0.239201
$a_{\perp A}=a_{\perp G}=0$	4.183142863	2	0.123492923
$a_{\perp J}=0$	27.88413	1	0
$a_{\perp A}=a_{\perp G}=a_{\perp J}=0$	34.85149313	3	0

A restricted version of the common trend is estimated by dropping Addis Ababa and Gonder and is given below. And it shows that both Bale Robe and Shashimiene are having more or less the same impact but the deficit market of Jimma is having the next stronger impact following both.

$$F_A = -10.66157741B - 8.422869842J + 10.11121128S$$

When the second combination is considered still producer centers are observed to have more say in price formation in the period of 1996 to 2003. The highest impact is coming from surplus market of Shashimiene followed by another surplus market of Bale Robe.

$$F_N = -3.2450884N - 9.2902753B - 6.6893742J - 4.7592718G + 11.280918S$$

Observing deficit markets the highest impact is coming from Jimma followed by Gonder. And the secondarily central market of Nazareth is having the weakest impact on the common integrating trend. Table 71 below is providing statistical justification about the conclusion that the important markets in the price formation or the determination of the common trend are Shashimiene, Bale Robe and Jimma.

Table 71 the statistical significance of impact of markets in the rule of one price or common trend on market combination which includes Addis Ababa

Null	L_P	DF	probability
$a_{\perp N}=0$	0.651329121	1	0.419637885
$a_{\perp G}=0$	1.00927	1	0.315078
$a_{\perp N}=a_{\perp G}=0$	2.311928531	2	0.314753884
$a_{\perp J}=0$	30.20327	1	0
$a_{\perp N}=a_{\perp G}=a_{\perp J}=0$	34.17580065	3	0

And given there are no lag of the first difference variables used in the VECN, when Gonder and Nazareth are dropped both combinations are having the same common trend.

$$F_N = -10.66157741B - 8.422869842J + 10.11121128S$$

So in both combinations price are discovered mainly in producer centers and Jimma and they are enforced in to other markets through the rule of one price which ever combination of market we take. Actually the six markets are following one common trend and the markets are rejected at six dimensions because normality was not attained in their error vector. So six markets made from two producer centers, two from deficit markets and two central markets are operating under rule of price. And even though the markets are observed to go in to disarray and confusion when shocks are initiated in surplus markets actually prices are discovered mainly in producer centers of Bale Robe and Shashimiene and customer center of Jimma.

6. <u>Implication for price stabilization effort</u>

Since the price of Shashimiene is weakly exogenous intervention could be very effective if it was targeting this market. However given the observed volatility cluster it could be very expensive to manage price stabilization effort in Shashimiene. In short price stabilization effort in Shashimiene will need only local information processing with out considering national dynamics but since shocks are coming in cluster it will need a higher logistic to make it effective. In Bale Robe the advantage is related to the fact that there is no volatility cluster which implies shocks are coming very randomly with out being clustered in to the same time. As result there is no need for large logistics to intervene in this market. The problem is that the market is not weakly exogenous means its prices are affected by dynamics and shocks initiated in other markets. So intervention in Bale Robe can't be done at local level with out processing the national dynamics of prices and their determinants. But one additional problem for both markets is that other markets will have low capacity to process information initiated in surplus markets and in short run they may go in to confusion. These can result on short run increase in volatility if price stabilization efforts are targeting producer centers but it will be sorted out in few months. The last option is to target the deficit market of Jimma. The advantage of this market is that it will have low impact on short run volatility as the markets are observed to having good processing capacity for information coming from deficit market and it will need lower logistic compared to Shashimiene. However there is need to consider national dynamics as the market is not weakly exogenous to shocks coming from other markets.

So if there is no problem of logistics to handle volatility cluster and if the objective is long run stabilization only the target market for stabilization have to be the surplus market of Shashimiene. And this can be done at local level with out national level analysis of economic parameters. If short run stabilization is as important as long run stabilization of prices there is need to focus on customer center of Jimma but with well developed analysis on national economic dynamics to make sense the pattern of national market. But the impact of Jimma will not as effective as either Shashimiene or Bale Robe. However if effective stabilization in long run is needed and there is effective institutional capacity to make analysis of the national economic dynamics it is better to select Bale Robe. And this can be done with lower logistics compared to Shashimiene but it will initiate increase in volatility in short run.

However given the existence of Ethiopian Grain Trade Enterprise (EGTE) to make detailed analysis in price dynamics it may not be very problematic to target markets which are not weakly exogenous. So it is logical if the interventions are targeting in Jimma and Bale Robe. Moreover with use of adequate logistics Shashimiene could be also very use full market for intervention in wheat market. However if there is no adequate logistics like finance and mainly storage facility it may be rational to target only Jimma and Bale Robe, in short run. Second there is need to create transparency in intervention mainly in intervening in surplus market of Bale Robe and Shashimiene; since the wheat market are having low capacity to process information coming from producer centers. This can reduce short run price volatility that can be resulted from intervention in surplus markets. And the direct implication to current policy is that the policy of targeting customer centers through subsidized distribution of wheat or injection of wheat in to central whole sale market of Addis Ababa and Nazareth is only solving the symptom than the source of the problem. These markets are not making the price but they are simply adjusting to price discovered in Shashimiene, Bale Robe and Jimma.

The two markets which are not in the common trend are Mekelle and Dire Dawa. Mekelle was observed to be pair wise cointegrated to Addis Ababa and Jimma but this was not extended in to higher dimension. And at third dimension the pair wise cointegration of Mekelle and Bale Robe it is extended to Nazareth but not to higher dimensions. These can be due to two different reasons. First it may be because Mekelle is cointegrated but is not part of the rule of one price. And this will be logical given the widely documented destabilizing impact of food aid in Ethiopian grain markets and given Mekelle is located in drought porn region; it may have its own impact on long run dynamics of prices. Second it could be because unrestricted trend is needed for Mekelle but not for others but it is not possible to make such specification in VECM. It would be preferable if there is additional model which allows restrict trend for Mekelle but unrestricted constant for other markets. So there is probability that Mekelle to be well cointegrated with other markets but may not be in the rule of one price. Dire Dawa is also observed to be cointegrated to both Bale Robe and Jimma but is not part of the one common trend under unrestricted constant model. So there is high probability that both Dire Dawa and Mekelle will be also affected by intervention in the other 6 markets even though they are not part of the one price system.

4. **Conclusion**

The most important conclusion of this paper is related to the identification of three markets where effective stabilization intervention can be done. These are the surplus market of Shashimiene and Bale Robe and deficit market of Jimma. Based on the reality faced by the policy makers each market also have its own advantage and disadvantage which can make it preferable site for intervention or not. Unfortunately the current policy of distributing wheat at subsidized price may be reducing the social cost of inflation; however it is not solving the source of the problem, but only the symptom.

Strength reversal is theoretically possible and apparently it was found in the data given the combination of 5 markets which include Nazareth will not be found by search

procedure used in earlier papers. However these is found to be apparent problem only than real since Bale Robe, Jimma, Gonder, Nazareth, Addis Ababa and Shashimiene are observed to follow one common trend in six dimension, even though their error vector was not normally distributed. This is related to limitation of the rank test which is developed for identically and independently distributed Gaussian errors, especially for moderately small sample sizes. Unless the errors are white noise rank test has weak power and have tendency to find cointegration when there is none. But because the errors are not white noise it does not mean there is no cointegration between the markets, only rank test can't be taken at face value. So there was cointegration between these six markets as proven by common trend they share, even though it can't be proved or disproved by rank test. So even though the search procedure followed by Gonzalo and Pitarakis (2000) is not theoretically sound it is observed to work for Ethiopian wheat markets.

Means the search can be simplified by building on those markets which are well cointegrated at lower dimension under one common trend by making the necessarily testes in each stage. But two things needs to be clear first these method is not theoretically sound but practically found to work in both Brazil rice market and Ethiopian wheat market. Second the order with highest normality may not be attained if markets are normalized against the markets which are having strong cointegration in lower dimension. So all possible permutation of markets have to be tested, when adding markets to already cointegrated markets under rule of one price.

However the assumption that the most important markets in higher dimension will be found to be cointegrated at lower dimension is not true since the most important market in Ethiopian wheat market is Shashimiene but it was not found to be cointegrated with other markets until 5th dimension.

Moreover in less developed economy with high level of market failure cointegration have to do with many factors including transportation convinces and distance but not with transportation convince and distance alone. The occurrence and depth of market failures in information, credit, storage, risk and other sides of the market can have impact on level of market integration. Moreover the development of complementary institutions to handle those market failures may also have impact on traders' capacity to create effective and efficient space utility. So the search four boundary, if there is any in under developed market, can't be distance based still routine search is needed at each dimension. And this is proven in the analysis.

Bibliography

1. Asfaw Negassa (1998) 'vertical and spatial integration of grain markets in Ethiopia: Implication for grain markets and food security polices' Grain Market Research Project working paper, No. 9
2. Bollerslev, T. (19860) 'Generalized Autoregressive Conditional Heteroskedasticity', Journal of Econometrics, Vol. 31, pp. 307-327.
3. Breusch, T S, (1978) "Testing for Autocorrelation in Dynamic Linear Models," Australian Economic Papers, Blackwell Publishing, Vol. 17, pp. 334-55,
4. Brüggermann, Ralf, Helmut Lütkepohl and Pentti Saikkonen (2006) 'Residual autocorrelation testing for vector error correction models' Journal of Econometrics Vol. 134, pp. 579–604
5. Cavaliere et al (2009) 'Co-integration Rank Testing under Conditional Heteroskedasticity' CREATES Research Paper, No. 22
6. Chao, John C. and Peter C. B. Phillips (1999) 'model selection in partially none stationery vector autoregressive process with reduced rank structure', Journal of Econometrics, Vol. 91, pp. 227-271
7. Davidson, Russell and James G. Mackinnon (1993) Estimation and Inference in Econometrics, Oxford University Press, Oxford – New York
8. Demiroglu, Ufuk (2000) 'Residual-Based Tests for Normality in Auto regressions: Asymptotic Theory and Simulation Evidence' Journal of Business & Economic Statistics, January (2000)
9. Dickey, D.A. and W.A. Fuller (1979) "Distribution of the Estimators for Autoregressive Time Series with a Unit Root," Journal of the American Statistical Association, Vol. 74, pp. 427–431
10. Eleni Z. Gabre-Madhin (2001) 'Market Institutions, Transaction Costs, and Social Capital in the Ethiopian Grain Market', IFPRI research report 124
11. Eleni Z. Gabre-Madhin, Wolday Amha, Ethiopis Tafara, John Schluter, Tilahun Teshome and Getinet Kilkile (2003) 'Getting Markets Right in Ethiopia: An Institutional and Legal Analysis of Grain and Coffee Marketing' IFPRI Final study report
12. Engle, R. (1982) Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation, Econometrica Vol. 50, pp. 987-1008.
13. Engle, Robert F and C. W. J. Granger (1987) 'Co-Integration and Error Correction: Representation, Estimation, and Testing', Econometrica, Vol. 55, pp. 251-276
14. Gebremeskel Dessalegn, T.S. Jayne and J.D. Shaffer (1998) 'Market structure, conduct and performance: constraints in performance of Ethiopian grain markets' Grain Market Research Project working paper, No. 8
15. Godfrey, Leslie G, (1978) "Testing for Higher Order Serial Correlation in Regression Equations When the Regressors Include Lagged Dependent Variables," Econometrica, Econometric Society, Vol. 46, pages 1303-10
16. Gonzalo, Jesus and Clive Granger (1995) 'Estimation of common long memory components in cointegrated systems', Journal of Business and Economic Statititics, Vol. 13, PP. 27-33
17. Gonzalo, Jesus and Jean-Yves Pitarakis (2000) 'Lag Length Estimation in Large Dimensional Systems' EconWPA, No. 0108003

18. Gonz´alez-Rivera, Gloria and Steven M. Helfand (2001) 'The Extent, Pattern, and Degree of Market Integration: A Multivariate Approach for the Brazilian Rice Market' American Journal of Agricultural Economics. Vol. 83, pp. 576–592

19. Greene, William H. (2003) <u>Econometric Analysis</u>, 5th edition, Prentice Hall Inc. New Jersey – USA

20. Gutiérrez, C. E. Carrasco, Reindldo C. Souza and Osmani T. de Carvalho Guillén (2007) 'Selection of optimal lag length in cointegrated VAR models with weak form of common cyclical features' Brazil central bank working paper series, No. 139

21. Hamilton, James D. (1994) <u>Time Series Analysis</u>, Princeton University Press, Princeton, New jersey

22. IMF (2008) 'The Federal Democratic Republic of Ethiopia: Selected Issues' IMF Country Report No. 08/259

23. Jarque, Carlos M.; Anil K. Bera (1980) "Efficient tests for normality, homoscedasticity and serial independence of regression residuals". Economics Letters, Vol. 6 pp. 255–259

24. Jarque, Carlos M.; Anil K. Bera (1981) "Efficient tests for normality, homoscedasticity and serial independence of regression residuals: Monte Carlo evidence". Economics Letters, Vol. 7, pp. 313–318

25. Johansen, S. (1988) "Statistical analysis of cointegration vectors" . Journal of Economic Dynamics and Control, Vol. 12, pp. 231–254.

26. Johansen, S. (1991) "estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models' Econometrica, Vol. 59 pp. 1551-80

27. Johansen, S. (1992) "Determination of cointegration rank in the presence of linear trend", Oxford Bulletin of Economics and Statistics, Vol. 54, pp. 383 - 97

28. Khim, Venus and Sen Liew (2004) 'Which Lag Length Selection Criteria Should We Employ?' Economics Bulletin, Vol. 3, pp. 1−9

29. Lee, Tae -Hwy and Yiuman Tse (1996) 'Cointegration tests with conditional heteroskedasticity' Journal of Econometrics, Vol. 73, pp. 401-410

30. Ng, Serena and Pierre Perron (2001) 'Lag length selection and the construction of unit root tetes with good size and power' Econometrica, Vol. 69, pp. 1519–1554

31. Pesaran, M. Hashem and Yongcheol Shin (1996) 'Cointegration and speed of convergence to equilibrium', Journal of Econometrics, Vol. 71, pp. 117- 143

32. Peterson, Kerry (2000) <u>An Introduction to Applied Econometrics: Time Series Approach</u>, Palgrave Macmillan – New York

33. Rashid, Shahidur (2004) 'spatial integration of maize markets in post-liberalized Uganda', Journal of African Economies, Vol. 13, PP. 102-133

34. Ravallion, M. (1986) 'Testing market integration', American Journal of Agricultural Economics, Vol. 68, pp. 102 –109

35. Sadoulet, Elisabeth and Alain de Janvry (1995) <u>Quantitative Development Policy Analysis</u>, The Johns Hopkins University Press, Baltimore and London

36. Ulimwengu, John M., Sindu Workneh and Zelekawork Paulos (2009) 'Impact of Soaring Food Price in Ethiopia: Does Location Matter?' IFPRI Discussion Paper 00846

37. Verbeek, Marno (2004) <u>A guide to modern econometrics</u>, 2nd edition, John Wiley & Sons Ltd Chiehester – England

38. Winker, Peter and Dietmar Maringer (2004) 'Optimal Lag Structure Selection in VEC Models' Computing in Economics and Finance, Vol. 155

39. World Bank (2008/9) <u>Africa development Indicators</u>, the International Bank for Reconstruction and Development/The World Bank, Washington, D.C.- U.S.A.