Applying cointegration to test PPP and UIP

March 6, 2020

In this practical illustration, we will investigate whether there is evidence of two prominent international parity relationships in South African data: the purchasing price parity relationship and the uncovered interest parity relationship. These are hypothesized equilibrium conditions on exchange rates that have been subject to much testing. Typically, the evidence is weak, especially for PPP.

1 Theory

1.1 Purchasing power parity

In international trade the strongest form of the Purchasing Power Parity hypothesis states that, once converted into the same currency via the nominal exchange rate E_t (domestic currency units per foreign currency unit), the aggregate foreign price index, P_t^*, and the aggregate domestic price index, P_t, should be identical. That is:

$$P_t = E_t P_t^*$$

Or equivalently, that the real exchange rate, Q_t is identically equal to one in all periods:

$$Q_t = \frac{P_t}{E_t P_t^*} = 1$$

In logs ($q_t = \ln Q_t$):

$$q_t = p_t - p_t^* - e_t = 0$$

This extreme version requires that each and every price satisfies the law of one price. We can only expect that to hold if all goods are equally easy to trade so that arbitrage ensures this result. The law on one price makes sense for the price of gold, but less so for hamburgers (you cannot arbitrage fresh burgers in New York and Stellenbosch directly, although the ingredients of fresh burgers may be more easily traded/arbitraged) Since different prices may adjust at different speeds, a weaker statement is that PPP holds in the long run - that is, that deviations from $q_t = 0$ eventually die out. i.e. that the logarithm of the real exchange rate reverts to its (unique) long run mean of 0 in the absence of shocks. A slightly stronger version would be that the variance and autocovariances of the real exchange rate are constant in the long run as well, or equivalently that the log of the real exchange rate q_t is a stationary process.

If p_t, p_t^* and e_t are $I(1)$ then PPP would imply they are cointegrated with cointegrating relationship $[1, -1, -1]$. In matrix form the cointegration hypothesis of PPP can be stated as:

$$\begin{bmatrix} 1 & -1 & -1 \end{bmatrix} \begin{bmatrix} p_t \\ p_t^* \\ e_t \end{bmatrix} = q_t \sim I(0)$$

Other real world issues may also affect this hypothesized relationship, e.g. risk premia may induce a permanent deviation from this relationship, or lead to a long run relationship with coefficients different from unity.

1.2 Uncovered Interest Parity

In international finance, the Fischer model argues that the exchange rate adjusts to interest rate differentials across countries. That is if the foreign interest rate is higher than the domestic interest rate, the expectation is that the exchange rate should depreciate (in the conventions here: e_t is expected to increase:

$$e_t - e_{t+1}^e = i_t^* - i_t$$

The spot exchange rate e_t is typically $I(1)$, so the expected exchange rate will also be $I(1)$. Economic theory strongly implies that $s_t = e_t - e_{t+1}^e$ should be stationary. Unfortunately, we do not have measures of exchange rate expectations. The next best thing would be to use the futures market: e.g. the three month future rand-dollar exchange rate should be an unbiased proxy for the three month ahead expected spot exchange rate. Unfortunately again, getting historical series on futures is not easy. But to test the theory, we can at least do the following: if i_t^* and i_t are both $I(1)$, then UIP implies that they should be cointegrated, such that $i_t^* - i_t$ is stationary. I.e. the left and right hand sides of the UIP equation are both stationary. We will treat the residual of the long run relationship, s_t, as the empirical measure of expected depreciation. In matrix form the cointegration hypothesis of PPP can be stated as:

$$\begin{bmatrix} 1 & -1 \end{bmatrix} \begin{bmatrix} i_t^* \\ i_t \end{bmatrix} = s_t \sim I(0)$$

Again, real world concerns may affect this relationship. Between South Africa and the rest of the world, a risk premium would allow a permanent interest rate differential that would not induce a change in the interest rate.

1.3 Joint testing

We can state both hypotheses simultaneously as follows:

$$\begin{bmatrix} 1 & -1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} p_t \\ p_t^* \\ e_t \\ i_t^* \\ i_t \end{bmatrix} = \begin{bmatrix} q_t \\ s_t \end{bmatrix} \sim I(1)$$

This will be our goal starting from a general case where all the coefficients in the cointegrating relationship are left unrestricted, B (to keep with coding conventions of both Eviews and R), and test a variety of restrictions to see if the parity relationships hold:

2 Empirical application to South Africa

In this practical we will do very simple analysis to see if there is evidence of these relationships, along the lines of:

Johansen, S. and Juselius, K., 1992. "Testing Structural Hypotheses in a Multivariate Cointegration Analysis of the PPP and the UIP for the UK" Journal of Econometrics 53(1-3) pp.211-244.

2.1 Data

I followed Johansen and Juselius (1992) as closely as I could. The data for this exercise are all taken from the extremely easy to use and free database collected and collated by the St Louis Federal Reserve (https://fred.stlouisfed.org/). The full range of the monthly data is from January 1981 to May 2019, although not all variables are available for the full time period, and we will use a restricted sample which we will motivate below.

Variable Name	Content
e_zar_usd	the logarithm of the average of the R:$ exchange rate in each month
e_zar_eur	the logarithm of the average of the R:€ exchange rate in each month
p_zar	the logarithm of the average of the SA producer price index in each month
p_usa	the logarithm of the average of the USA producer price index in each month
p_eur	the logarithm of the average of the Euro area producer price index in each month
i_zar	3-Month or 90-day Rates and Yields: Interbank Rates for South Africa
i_eur	3-Month or 90-day Rates and Yields: Interbank Rates for the Euro area
i_usa	3-Month or 90-day Rates and Yields: Interbank Rates for the United States

2.2 Data range

One's first intuition is that one should use all data available. However, this agnostic approach ignores potential structural breaks that may obscure the question at issue. The major concern in this case is that after the 2008 Financial crisis, interest rates hit the zero lower bound in the USA and the Euro area, but not in South Africa. This could arguably change the equilibrium relationship between these interest rates for that period. Other obviously important structural changes are (1) the Euro was only an effective monetary (and thus banking) union since the mid 1990's, (2) the digitalization of the financial economy became progressively more pervasive since the 1980's, and (3) South Africa's was excluded from international markets before 1994.

We can heuristically investigate these issues by viewing joint graphs of the relevant data. Let's look at full sample graphs for:

- exchange rates
- interest rates
- command: Open as group: select variables > right-click > open as group. Then View > Graph

For this reason, we will restrict our estimation sample to range from January 1995 to June 2007. Did I cherry pick these dates to get the results I want? That is up to you to argue and evaluate the validity of.

- set sample to: 1995m01 2007m06

2.3 Check order of integration of data

To conduct a cointegration study, all variables must be $I(1)$ so the first step is to test for unit roots using the most appropriate test statistic. This includes the question of whether we should allow for deterministic parts (a constant and/or a trend).

We will take the following steps for each series:

- graph to determine the appropriate deterministic components of the test,
- conduct the appropriate tests, but check the other options as well.
  - Augmented Dickey Fuller
  - KPSS
- Commands: open variable
  - graph: View > Graph
  - unit root test: View > Unit Root Tests > Standard Unit Root Tests > select test and specification. First in levels then in differences

3 Testing for UIP alone

Let us start with the simplest type of cointegration study: a bivariate case where there can be at most 1 cointegrating relationship. We proceed in the following steps

3.1 Estimate unrestricted VAR, reduce and evaluate adequacy

We always start with an unrestricted VAR. If there is cointegration,

- The residuals of the estimated unrestricted VAR must be stationary
- For the VAR to be a valid representation of the data generating process, the residuals must at least be white noise
- If we are using Maximum Likelihood methods based on Gaussian errors, the errors should also test as normal (although the Johansen method is somewhat robust to certain type of deviations from normality)
- Since we know our variables are I(1), we now expect (and require) that there be unit roots in the system. Since this is a statistical estimate, we will only by chance find roots exactly equal to 1. Later we will impose this explicitly.

Estimate the unrestricted VAR

- Select the variables, right-click, and select "Open as VAR"
  - start with a generous lag structure, e.g. 12 lags (one year) for a monthly variables. Since financial markets should move quickly to wipe out arbitrage opportunities, we suspect that this will be an over-parameterized model.
  - There is no obvious trend in the data, so we stick to just constants in the specification.

Reduce the VAR to be parsimonious

- Check lag exclusion tests and lag-length criteria. Commands: View > Lag Structure
- Both suggest that a VAR(2) is probably appropriate

Estimate a VAR(2) and evaluate the adequacy:

- Inspect the characteristic roots of the estimated process: View > Lag Structure > AR Roots Table and/or Graph
- Inspect the residuals visually (command button: Resids)
  - Note that there are some periods with large residuals. This will almost certainly make the residuals test as non-normal. What to do about this is not uncontroversial, but we can never expect a simple two variable equation to model all possible economic forces, and for the purposes of this exercise we are interested in the "normal times" cointegration between these interest rates.

- As in Johansen in Juselius, it may be necessary to generate dummy variables to remove these "outliers". If one does this, one must be absolutely candid about ones choices and motivate them by real world events. In this case large financial distruptions might be to blame for these outliers:
  - Dot-com bubble reached its peak in the USA in March 2000, with the crash following in early 2001, followed by the 9/11 attacks
  - In 1998, Russia defaulted on its debt which sharply increased yields and all other interest rates in developing markets. This disturbance lasted very briefly in the South African market (as can be seen by the simple plot of the interest rate), hence both the spike and correction are likely to cause large outliers
- We will first run a battery of tests on the residuals, then generate dummies, re-estimate and re-run the tests to see how they change

- Test the residuals. View > Residual Tests
  - Autocorrelation.
    - view cross-correllogram
    - autocorrelation tests
    - broadly, there is no strong evidence of strong autocorrelation in the residuals
  - Heteroscedasticity
    - One of the series seem to have some autocorrelation in the squared residuals. This could be due to the outliers, so let's see if adding dummy variables for the extreme events reduces the problem
  - Normality
    - Neither residual series seems very close to normal, so we will have to be cautious with any maximum likelihood based test
- Generate dummies for the largest outliers in each series and re-run the tests:
  - We will generate dummes for: 1998m6 1998m9 1999m10 2001m1 and 2001m9
  - The command is: series dummy1998m6 = @recode(@date=@dateval("1998m6"),1,0)
  - Note that we are effectively reducing our sample by 5 periods: the dates we "dummy-out" have no influence on the fit of the dynamics of the model

3.2 Test for cointegration, number of cointegrating relationships

If the unrestricted VAR is adequate (which we assume for now), one can move on to test for cointegration which is based on the unrestricted VAR estimate. View > Cointegration Test

- There are 5 different possible specifications for the deterministic parts.
  - We note a permanent difference in the two interest rates - the South African interest rate is always higher than the US interest rate. Thus a restricted intercept in the cointegrating equation is definitely sensible. This would represent a (constant) risk premium in the equilibrium.
  - There also is a general decrease in interest rates over time
  - Simplest option: test all possible cases at the same time. Caution: the final one chosen should be motivated by economic arguments, not desirable outcomes!
  - The tests suggest three specifications where there may be cointegration by at least one of the Johansen tests
  - The most economically reasonable one is an intercept in the cointegrating relationship, but no trend. We will also briefly investigate the linear trend case

3.3 Estimate the restricted VECM model

Once the number of cointegrating relationships has been established, we can estimate the VECM restricted, which imposes the number of cointegrating relationships found by the tests (and/or suggested by economic theory). Estimate, select "Vector Error Correction" under the "VAR type" option, click on the "Cointegration" tab and select the appropriate option

- In principle, we should redo all our specification tests to ensure that the restriction we imposed (1 cointegrating relationship) is still congruent with the data generating process, but for illustrative purposes, we will skip that and carry on
- Inspect the estimation output for reasonability
  - Signs of the long run relationship
  - Responses of the dynamics of the variables - does it go in the right direction?
  - Roots of the estimated process
  - Cointegration relationship (long run error)

3.4 Test economic hypothesis as further restrictions on the estimated system

If we are satisfied with the restricted VECM, we can now turn to economic hypothesis testing by imposing further restrictions on the estimation

- Which of the variables adjusts to disequilibrium? Given the disparity in the sizes of the economies, it would be very surprising if the US market adjusted to deviations. It is quite reasonable to expect the SA market to respond to deviations.
  - Estimate, select the "VEC Restrictions" tab
  - Note: if you impose restrictions, you have to manually impose the normalization
  - To test for weak exogeneity of the US interest rate: B(1,1) = 1, A(1,1)=0
  - Weak exogeneity of the US interest is not rejected by the data - so we can use an ARDL approach.
- Does the strict version of the UIP hold?
  - Restriction (with the weak exogeneity restriction): B(1,1) = 1, A(1,1)=0, B(1,2)=-1

4 Testing for PPP and UIP jointly

In this situation we have 5 $I(1)$ variables, which allows up to 4 cointegrating relationships. Economic theory suggests that there should be two long run relationships between these variables. We will investigate whether such a restriction is congruent with the data generating process of the variables. We will follow exactly the same steps as in the section above, so I present them more briefly here.

4.1 Estimate unrestricted VAR, reduce and evaluate adequacy

We always start with an unrestricted VAR. If there is cointegration,

- The residuals must at least be white noise
- If the residuals are sharply non-normal, one must be cautious and do some careful reading of the literature about how such deviations will affect test sizes, power and validity
- The process should have some (very near) unit roots.

Estimate the unrestricted VAR

- Select the variables, right-click, and select "Open as VAR"
  - To make sure your system is estimated in the same order as mine (and as in the analytical order in the theory above) make sure that you select them in the correct sequence (Hold Ctrl and select in sequence: p_zar, p_usa, e_zar_usd, i_usa, i_zar)

    * You can check the sequence in the "Endogenous variables" box in the VAR Specification window
    * The sequence does not matter directly, only to keep track of which coefficients we need to restrict to get the empirical version of the two parity relationships
  - start with a generous lag structure, e.g. 6 lags (one year) for a monthly variables. Since we now have price series that move much more slowly, a more generous lag structure might be necessary.

Reduce the VAR to be parsimonious

- Check lag exclusion tests and lag-length criteria. Commands: View > Lag Structure
- Both suggest that a VAR(2) is probably appropriate

Estimate a VAR(2) and evaluate the adequacy:

- Inspect the characteristic roots of the estimated process: View > Lag Structure > AR Roots Table and/or Graph
- Inspect the residuals visually (command button: Resids)
  - Note that there are some periods with large residuals. This will almost certainly make the residuals test as non-normal. What to do about this is not uncontroversial, but we can never expect a simple two variable equation to model all possible economic forces, and for the purposes of this exercise we are interested in the "normal times" cointegration between these interest rates.
  - The residuals from the p_zar series, however, look problematic
  - We will first run a battery of tests on the residuals, then generate dummies, re-estimate and re-run the tests to see how they change
- Test the residuals. View > Residual Tests
  - Autocorrelation.
    * view cross-correllogram
    * autocorrelation tests
    * broadly, there is no strong evidence of strong autocorrelation in the residuals
  - Heteroscedasticity
    * One of the series seem to have some autocorrelation in the squared residuals. This could be due to the outliers, so let's see if adding dummy variables for the extreme events reduces the problem
  - Normality

        * Neither residual series seems very close to normal, so we will have to be cautious with any maximum likelihood based test
    - The VAR(2) does not seem to be sufficient to capture enough of the time-series dependence

- Estimate a VAR(4) and repeat tests
    - This looks much better, although not perfect. We will continue with this specification
    - The graph of the residuals suggest outliers at: 1998m7 , 2001m12 (e_zar_usd resid), 1999m10, 2001m1 (i_usa), 1998m6, 1998m10 (i_zar)
    - Generate dummies for the largest outliers in each series and re-run the tests:

4.2 Test for cointegration, number of cointegrating relationships

If the unrestricted VAR is adequate (which we assume for now), one can move on to test for cointegration which is based on the unrestricted VAR estimate. View > Cointegration Test

- There is a trend in the price data, but we don't expect a trend in the cointegrating relationship so we will use this as our best gues of the process
- Technically the Johansen Test does is not derived with exogenous variables, but even Johansen and Juselius do this
- At 5% there is evidence for 2 cointegrating relationships
    - This is very sensitive to the dummy variables, which is not ideal

4.3 Estimate the restricted VECM model

Once the number of cointegrating relationships has been established, we can estimate the VECM restricted, which imposes the number of cointegrating relationships found by the tests (and/or suggested by economic theory). Estimate, select "Vector Error Correction" under the "VAR type" option, click on the "Cointegration" tab and select the appropriate option

- In principle, we should redo all our specification tests to ensure that the restriction we imposed (2 cointegrating relationships) is still congruent with the data generating process, but for illustrative purposes, we will skip that and carry on
- Inspect the estimation output for reasonability
    - Signs of the long run relationship

  - Responses of the dynamics of the variables - does it go in the right direction?
  - Roots of the estimated process
  - Cointegration relationship (long run error)

- The interpretation is now difficult in economic terms as the estimation output just gave an arbitrary normalization.

4.4 Test economic hypothesis as further restrictions on the estimated system

If we are satisfied with the restricted VECM, we can now turn to economic hypothesis testing by imposing further restrictions on the estimation

- The weakest form of our PPP and UIP hypotheses can be written as:

$$\begin{bmatrix} B_{11}=1 & B_{12} & B_{13} & B_{14}=0 & B_{15}=0 \\ B_{21}=0 & B_{22}=0 & B_{23}=0 & B_{24}=1 & B_{25} \end{bmatrix}$$

- The test for the hypotheses even in their weakest form seem only to be accepted if the right set of dummy variables are included. This is evidence of a pretty unstable result.
  - The US is not the largest trading partner of SA, so we might expect the evidence for PPP to be weak.
  - The EU is a much bigger trade partner, so we can do the same analysis with them
  - Or we can consider a "composite" trading partner by using trade weighted data with all our biggest trading partners (this is actually what Johansen and Juselius did)
    * But more complications arise: the UIP is a financial market concept - should we use different weighting schemes relating to our most important capital market links?