A Vector Error Correction Modelling of the Composition of Government Spending and the Real Exchange Rate in Nigeria

Victor E. Oriavwote

Department of Arts and Social Sciences
Delta State Polytechnic, Otefe-Oghara
E-mail: oriavwote40@yahoo.com

Ben U. Omojimite

(Corresponding author)
Department of Economics, Delta State University, Abraka
Email: buomojimite@yahoo.com

Doi:10.5901/mjss.2012.v3n2.543

Abstract: The main objective of this study has been to empirically investigate the relationship between government expenditure and the Real Exchange Rate in Nigeria . Using data covering 1970 to 2010, the Vector Error Correction and Cointegration tests results established a long run relationship among the components of government expenditure and the Real Exchange Rate. The result showed that government investment has delivered more productivity gains in the non-tradable sector than in the tradable sector, while an increase in government consumption increased the relative demand for non-tradable. The government should thus shift attention to the expansion of the tradable sector.

Key words: Government Consumption, Government Investment, Real Exchange Rate and Vector Error Correction

1. Introduction

The Real Exchange Rate (RER) is a significant factor in the development process of an economy as both its level and stability are important in increasing exports and private investment. While misalignment of the RER can be an undervaluation or overvaluation, overvaluation has been the common case in developing countries (Edwards, 1989). An overvalued RER reduces profits in the tradable goods sector, thereby reducing investment in the sector. This has a negative effect on export and the trade balance. Persistent overvaluation of the Real Effective Exchange Rate (REER) may also lead to currency crisis (Xiaopu, 2002). The increasing overvaluation of the exchange rate that began in Sub-Sahara Africa in the early 1980s contributed to the poor performance of the current account balance in the region (Ghura and Grennes, 1993). Thus as noted by Mungule, 2004), the RER does not only affect the general performance and international competitiveness, but also different sectors of the economy including foreign trade flows, balance of payments, external debt crisis, employment, structure of production, consumption and allocation of resources in the economy.

In Nigeria, the RER has been an important player in the growth of the economy. Given the import dependent nature of the Nigerian industrial sector, the continued depreciation of the naira exchange rate vis-à-vis the currencies of other major trading partners meant that more resources would be needed to increase domestic output. A depreciating Exchange Rate in the absence of domestic sources of inputs coupled with inadequate infrastructure has raised the cost of production in Nigeria, which in turn made locally produced goods less competitive compared to imported goods. This has reduced the benefits of cheaper exports expected from the depreciation of the currency. Similarly, the overdependence of the economy on imported capital goods implies that a depreciating exchange rate would crowd out the marginal investment as a result of high investment cost (Nnanna, Englama and Odoko, 2004). This depreciation of the naira exchange rate in

Nigeria has resulted in a shift in finance to the tradable sub-sector. In his view, Obadan (2006) identified the factors that lead to instability in the RER in Nigeria to include amongst others round-tripping, weak production base, import dependent production structure, weak non-oil export earnings, expansionary fiscal and monetary policies, fluctuations in crude oil earnings, unguided trade liberalization (which include removal in diesel and kerosene subsidies or reduction of subsidy in Petroleum Motor Spirit). In addition, the continuous depreciation of the naira has resulted in a monumental increase in the naira cost of executing government projects. This has led to a significant funding gap and sometimes inability to complete them. To be able to fund its projects, government has often resorted to expansionary monetary and fiscal measures that further put pressure on the naira exchange rate.

The question that comes to mind here is: what is the relationship between the composition of government spending and the RER in Nigeria? The objective of this paper is thus to assess the relationship between government expenditure and the RER in Nigeria using the Vector Error Correction Model (VECM). Specifically, the differential impact of both government investment and government consumption in Nigeria will be assessed in this paper. This distinction is important because our aim is to point out that government consumption and government investment may be expected to have different impacts on the RER in Nigeria. While an increase in government consumption is typically modelled as increasing the relative demand for nontradeable and thereby leading to real appreciation, a long run increase in public investment that delivers productivity gain in the tradables sector may generate real appreciation through the so-called Balassa-Samuelson mechanism. If public investment disproportionately raises productivity in the nontradables sector, it may lead to real deprecation. However, if productivity is increased in both sectors there is no long-run impact on the relative price of non-tradable and the RER (Vahagn and Philip, 2008). Other than this introductory section, the rest of the paper is divided into four sections. The second section is on the theoretical framework. This is followed by the third section which reviews empirical literature. The fourth section is on the econometric procedure and the fifth section concludes the paper.

2. Theoretical Issues

This section encompasses an adapted version of the standard two-sector small open economy model (Obstfeld and Rogoff, 1996 and Vahagn and Philip, 2008). The production functions for traded and non-traded goods are respectively.

$$Y_T = A^*_T F(L_T, K_T) = (A_T Z^{\alpha Z}) L_T^{\alpha L} K_T^{\alpha T} \quad (1)$$

$$Y_N = A^*_N G(L_N, K_N) = (A_N Z^{\beta Z}) L_N^{\beta L} K_N^{\beta T} \quad (2)$$

Where L and K stand for labor and capital, while Z stands for the public capital stock. It was assumed that total factor productivity in each sector is a composite of a sector- specific term (A^*_T, A^*_N) and the level of public capital. Accordingly, productivity in both sectors is enhanced by a larger stock of public capital but the impact was allowed to be potentially different across sectors $\alpha z \neq \beta z$). It was assume that $\alpha L + \alpha K = 1$, but $\beta L + \beta K < 1$. A fixed factor of production (normalized to 1) was incorporated in the non-traded sector such that the production function in that sector exhibits diminishing returns to labour and capital. The price of the traded good is equal to world price of the good and is normalized to 1, while the price of non-traded goods is P_N.

The accumulation functions for the private capital stocks in the traded and non-traded sectors are given by

$$\Delta K_T = 1^{K_T} - \delta K_T \quad (3)$$

$$\Delta K_N = 1^{K_N} - \Delta k_N \quad (4)$$

Where 1 denotes the level of gross investment and δ is the depreciation rate. The public capital stock evolves according to

$$\Delta z = I^Z - \delta Z \quad (5)$$

It was assumed that private capital formation in the traded and non-traded sectors only requires the traded good as an input, while public capital formation uses only the non-traded good as an input. The representative household has an instantaneous utility.

$$C = \frac{C_T^{1-\Upsilon} C_N^{\Upsilon}}{(1-\Upsilon)^{1-y} \Upsilon^{\Upsilon}} \quad (6)$$

The implication here is that optimal household expenditure shares on traded and non-traded goods are fixed at (1- Υ) and Υ respectively, with a unit elasticity of the relative consumption of non-tradables in relation to the relative price of non-tradeables.

The welfare-based price index consistent with equation (6) is

$$P = P_N^{\Upsilon} \quad (7)$$

It was also assumed that the price of the non-traded good in the rest of the world is fixed and normalized to 1, such that changes in P correspond to changes in the RER.

By assumption, the government runs a balanced budget, levying lump-sum taxes equal to the value of total government consumption and government investment.

$$T = G_T + P_N(G_N + I^Z) \quad (8)$$

Where G_T, G_N are the levels of public consumption of the traded and non-traded goods respectively and I^Z is the level of public investment.

Households own the domestic stocks of capital in the traded and non-traded sectors. There are no inter-sectoral or international capital adjustment costs, so that the return on capital is equal to the exogenously fixed world interest rate. In addition, households own the fixed factor in the non-traded sector and so receive the income accruing to that factor (the residual claimant on profits in the non-traded sector). Accordingly, households face the following budget constraint

$$AB = RB + r(K_T,K_N) + w(L_T, L_N - (I_N^K + I_T^K) - C_T - P_N C_N + \Pi_N - T \quad (9)$$

Where β is an international bond that pays the fixed real world interest r (in terms of tradables). $\Pi_N = (1 - \beta L - \beta K)P_N Y_n$ is the aggregate profit in the non-traded sector and T is the tax burden.

For simplicity, it is assumed that the aggregate labour supply is inelastic. Labor is perfectly inter-sectorally mobile, such that the equilibrium in the labor market is

$$L_N + L_T = L \quad (10)$$

The equilibrium in the non-traded goods sector is

$$Y_N = C_N + G_N + I^Z \quad (11)$$

While the trade balance is determined by

$$T_B = Y_T - C_T - G_T - (I_N^K + I_T^K) \quad (12)$$

Equations (1) to (12) together with the first-order conditions for private consumption and private investment and the profits of the non-traded sector form the system.

The primary interest is in the long-run behavior of the RER. Accordingly, the steady-state solution of the model was given attention. In order to obtain an analytical solution, it was assumed that there is no

depreciation. A benchmark steady state in which the levels of net foreign assets and government consumption are zero (in order to obtain a closed – form solution) was analyzed. The system was log-linearized around this benchmark, in order to examine the sensitivity of the steady-state RER to shifts in the steady state values of the exogenous variables.
In the benchmark steady state, the relative price of non-traded goods is

$$P_N = \left[\Pi L_N \frac{1-B_L-B_K\alpha_Z}{1-B_L} \frac{}{Z^{1-\alpha K}} \frac{\beta_Z}{1-\beta_k}^{1-\beta k} \right] \quad (13)$$

In the next stage, the system was log-linearized around this steady state and a solution was solved for. The equation of primary interest is the one governing the RER, $\bar{p} = \Upsilon\ \bar{p}_N$, with the relative price of non-traded goods given by

$$\bar{P}_N = -\ \bar{A}_N + \frac{1-\beta_K}{1-\alpha_K}\ \bar{A}_T + \mu_0 (r\ \bar{B} + [G_N - G_T] + \mu_1\ \bar{Z}) \quad (14)$$

Where hatted variables denote percentage deviations from the steady-state values. Equation (14) shows that an improvement in productivity in the non-traded sector generates real depreciation and a decline in the relative price of non-tradables, while an increase in productivity in the traded sector generates real appreciation and an increase in the relative price of non-tradable, where these forces operate according to the classic Balassa Samuelson mechanism. The other key coefficients are

and

$$\mu_O = \frac{\alpha_L (1-B_L-B_K)(1-\Upsilon)}{\alpha_L(1-\Upsilon)+B_L\Upsilon} > O \quad (15)$$

$$\frac{(1-B_K)\alpha_Z-(1-\alpha_K)\beta_Z}{\alpha_L} \quad (16)$$

Since $\mu_0 > 0$, the level and composition of spending matters for the RER. In particular, a country that is a long-run creditor ($\bar{\beta} > 0$) experiences real appreciation, since the interest income on the creditor position enables an increase in the steady state level of consumption. In the traded sector, this translates into a long-run trade deficit ($TB = -\ r\ \bar{\beta}$), while the increase in demand for the non-traded good generates real appreciation due to the presence of the fixed factor in the non-traded sector. Moreover, an increase in government consumption that is concentrated on non-traded goods ($[G_N - G_T] > 0$) also generates real appreciation by shifting the composition of aggregate consumption towards the non-traded sector.

Finally, the effect of an increase in the public capital stock on the RER is given by the coefficient p, which has an ambiguous sign. If an increase in public capital has a symmetric impact on productivity in both sectors ($\alpha_Z = \beta_Z$) and both sectors have similar capital shares ($\alpha_K = \beta_K$), the RER is unaffected by the level of the public capital stock. If $\alpha_Z = \beta_Z$ but the non-traded sector is less capital intensive ($\alpha_K > \beta_K$), then an increase in public capital generates real appreciation, by the same logic as a symmetric improvement in the sector specific productivity terms A_T and A_N. However, even if $\alpha_K > \beta_K$, it is possible to construct scenarios in which an increase in the public capital stock generates real depreciation if productivity in the non-traded sector is

more sensitive to the level of public capital than is productivity in the traded sector (($\alpha_Z = \beta_Z$). accordingly, the sign of the relation between public investment and the RER is ultimately an empirical matter.

3. Empirical Literature

Edwards (1989) pioneered the fundamental models of the determination of RER for developing countries. Edwards started by developing a theoretical model of the RER determination and then estimated its equilibrium value for panel of 12 developing countries. He used the conventional cointegration tests to analyze the relative importance of real and nominal variables in the process of RER determination in the short and long run. The study found that in the long run only real variables affect the long run equilibrium RER. In the short run, however, RER variability was explained by both real and nominal actors. More precisely, the most important factors identified in this study as affecting the equilibrium RER are the terms of trade and composition of government spending, the control of foreign exchange and the movement of goods (openness), technical progress and capital inflow. An increase in government consumption, capital inflows, terms of trade and a decrease in technological progress and openness appreciated the RER.

Ghura and Grennes (1993) used a panel of Sub-Saharan countries, to investigate the determinants of the RER and the impact of RER misalignment on economic performance. They employed a classical regression methodology and found that the RER appreciated with (i) an improvement in the terms of trade, (ii) a capital inflow, (iii) a decrease in openness, (iv) an increase in excess of domestic credit, and (v) an improvement in technology. The Study also found that nominal devaluation depreciated the RER. With regard to the impact of RER misalignment and variability, it found that RER misalignment and variability negatively affected income growth, exports, imports, investment and savings.

Elbadawi (1994) developed a model of the determination of the long run equilibrium RER. The fundamental determinants of the long run equilibrium RER in this model included the terms of trade, openness (a proxy for commercial policy), the level of net capital flows relative to GDP, the share of government spending in GDP and the rate of growth of exports (a productivity measure). Elbadawi empirically estimated his model on annual data for Chile, Ghana and India. The findings of this study suggested that, in all three countries, the RER and all the fundamentals identified in the model were non-stationary and cointegrated.

MacDonald (1998) presented a reduced form model of the RER to re-examine the determinants of RER in a long run setting. His model features productivity differentials, terms of trade effects, fiscal balances, net foreign assets and real interest rate differentials as key fundamental determinants of the RER. Using multivariate cointegration methods, he found evidence of a significant and sensible long run relationship, indicating that the fundamentals mentioned above have an important and significant bearing on the determination of both long and short run RER. Mkenda (2001) analyzed the main determinants of the RER in Zambia. The study presented an illustrative model based on the three-good production structure and employed cointegration analysis in estimating the long run determinants of the RER for imports and exports, and of the internal RER. The results of this study provided evidence that (i) a decline in the terms of trade and government consumption depreciated the RER for imports, while an increase in investment share of GDP appreciated the RER for imports; (ii) a decrease in the terms of trade, an increase in central bank reserves and trade taxes appreciated the RER for exports in the long run; (iii) in the long run, the internal RER was strengthened by a decrease in the terms of trade, an increase in investment share and the rate of growth of real GDP (A proxy for technological progress); (iv) in the short run, however, aid and openness depreciated the RER indices.

Joyce and Kamas (2003) re-investigated the factors that determined the RER in Argentina, Colombia and Mexico, distinguishing between real and nominal determinants. The study employed cointegration analysis, variance decompositions and impulse response analysis. Cointegration results established that the RER has an equilibrium relationship with real variables (terms of trade, capital flows, productivity and government

share of GDP). It excluded nominal variables (nominal exchange rate and money) and central bank intervention. In addition, an increase in all the real variables in their model appreciated the RER,The variance decomposition showed that the terms of trade and productivity, among other real variables, explained much of the variation in the RER. In the short run, however, the nominal exchange rate accounted for most of the variation in the RER of all three countries. Finally, the impulse response analysis revealed that shocks to nominal variables have only transitory effects on the RER, thus consistent with theoretical predictions.

Coricelli and Jazbec (2004) analyzed the phenomenon of RER appreciation that has characterized 19 transitional economies, which included a group of nine Central and Eastern European countries, three Baltic countries and seven former Soviet Union Countries. They used Ordinary Least Squares (OLS) regression to show that the RER is affected by the adverse initial conditions and structural reforms only in the first five years of the transition process. After this period, their results provided evidence that productivity differential, the share of non-tradable consumption in total private consumption and real government consumption negatively affected the RER, thus contributing to the real appreciation of the currencies of these economies. However, the Balassa-Samuelson effect (productivity differentials) seemed to dominate the determination of the RER in this study.

Mungule(2004) investigated the determinants of RER in Zambia. He used the RER as a function of terms of trade, capital inflow, closeness of the economy and excess supply of domestic credit. Using the cointegration technique, he discovered that the REER and the fundamental determinants have a long run equilibrium relationship. Vahagn and Philip (2008) investigated the composition of government spending and the RER. They developed a two sector open economy model in which an increase in government consumption is associated with real appreciation. Using panel data covering 1980 to 2004 and the cointegration technique, their result showed that government consumption typically leads to real appreciation. They also found that government investment has no significant impact on the RER for the set of EMU member countries.

4. Econometric Procedure

VAR Modelling and the Cointegration Approach

Vector autoregression (VAR) modelling and the cointegration approach provide not only an estimation methodology but also explicit procedures for testing the long-run relationship among variables suggested by economic theory.

According to the Granger Representation Theorem (Engle and Granger, 1987), if a P*1 vector, X_t, generated by (1-L) X_t = d + c(L) e_t, is cointegrated, then there exists a vector auto regression (VAR), an error correction, as well as a moving average (MA) representation of X_t. A set of variables X_t, which is cointegrated, refers to the existence of long-run equilibrium relationships among economic variables (Mungule, 2004). That is, though each series may be non-stationary, there may be stationary linear combinations of the variables. The basic idea is that individual economic time series variables wander considerably, but certain linear combinations of the series do not move too far apart from each other. In economic term, there is a long-run relationship among the variables.

The most common test for cointegration is the two-step procedure of Engle and Granger (1987) which performs well for univariate tests. The first step is to fit the cointegration regression, an ordinary least squares (OLS) estimation of the static model. The second step is to conduct a unit root test on the estimated residuals. To test for cointegration is just to test for the presence of a unit root in the residuals of the cointegrating regression. If the null of a unit root is rejected, then cointegration exists. However, the long-run parameter of the cointegrating vector estimated from this approach can be severely biased in finite samples. An improved procedure of cointegration test is that which allows for more than one cointegrating vector, as suggested in Johansen (1998) and Johansen and Juselius (1990).

Following Johansen and Juselius (1990), let the p variables under scrutiny follow a vector autoregression of order p (VAR(p)) as below.

$$X_t \; c + P_1X_{t-1} + \ldots + PpX_{t-p} + e_t \qquad (1)$$

Where X_t = n x 1 vector of economic variables in the model; c = n x 1 vector of constants or drift terms are innovations of this process and are assumed to be drawn from p- dimensional independently, identically distributed (i.i.d) Gaussian distributions with covariance G; and X_{p+1} ,. .. X_0 are fixed.
Where;

Pi = nxn matrixes of time invariant coefficients, I 1, . . .p, and
e = nx1 vector of i.i.d. errors with a positive covariance matrix.

Let Δ represent the first difference filter. The equation can be reparameterised into the equivalent form presented below

$$\Delta X_t = c + PX_{t-p} + \sum_{I-0}^{P-1} \tau_i \Delta X_{t-i} + \tau_t \qquad (2)$$

$$\text{Where } \tau_t = -\tau + \sum_{J-1}^{i} P_j, \quad \text{for, } I = 1, \ldots, P-1\text{-}, \tau + \sum_{i}^{p} P_j$$

The coefficient matrix P contains information about the long-run relationships among variables. Since e_t is stationary, the number of ranks for matrix P determines how many linear combination of X_t are stationary. If 0 < Rank (P) = r<p, there exists r cointegrating vectors that make the linear combinations of X_t to become stationary. In that case, P can be factored as "a" and "b", with "a" and "b" being matrixes. Here "b" is a cointegrating vector that has the property that bX_t is stationary even though X_t itself is non-stationary and "a" then contains the adjustment parameters.

Based on an unrestricted estimation that is parameterized in terms of levels and differences, Johansen (1988) proposed likelihood ratio statistics for testing the number of cointegrating vectors. First we must solve the eigenvalues of $|^eS_{pp} - S_{po}S_{00}^{-1}S_{op}| = 0$, where S_{00} is the moment matrix of the residuals from the ordinary least squares (OLS) regression of DX_ton ΔX_{t-1}. . . . ΔX_{t-p1}; S_{pp} is the residual moment matrix from the OLS regression of ΔX_t p on ΔX_{t-1} …… $\Delta X_{t-p+1+;}$ and S_{OP} is the cross product moment matrix. The cointegrating vector, b, is solved out as the eigenvectors associated with the r largest statistically significant eigenvalues derived using two test statistics, "maximum eigenvalue statistics" and "trace statistics". The first statistic tests the hypothesis that there are r=s cointegrating vectors against the alternative of r = s + 1 by calculating the maximum likelihood test statistics as $-T \ln(1-1_{s+1})$, where T is the sample size and 1_{s+1} is an estimated eigenvalue. The second statistic tests the hypothesis that there exists at most, r cointegrating vectors. If the test is performed by calculating trace statistics.

$$-T \sum_{I=}^{p} \ln\{ 1- \lambda i *)/ (1 - \lambda i)\} + 1$$

Where i* are eigenvalues obtained from cointegration analysis assuming there is no linear trend.
Thus, a proper assessment of the composition ofgovernment spending and the RER necessitated the estimation of the following equation:

$$LREER = b_0 + b_1LGCONS + b_2LGINV + b_3PRODUCTIVITY + U_t$$

Where

REER	=	Real Effective Exchange Rate
GCONS	=	Government Consumption
GINV	=	Government Investment
PRODUCTIVITY	=	Productivity is represented by the growth rate of Gross Domestic Product
L	=	Natural Logarithm
U_t	=	Random Variable

The results of the descriptive statistic for the variables is shown in table 1 below

Table 1: Summary of Descriptive Statistic

	GCONS	GINV	REER	PRODUCTIVITY
Mean	551618.8	283703.8	335.4242	0.482574
Median	34388.90	24775.50	130.2300	0.062191
Minimum	4162627	2272760	1075.780	9.375315
Minimum	143.6000	993.0000	57.47000	-0.891466
Std. Dev.	965888.4	554650.3	320.9957	1.719859
Skewness	2.178414	2.316480	0.973653	4.226820
Kurtosis	7.388690	3.239988	2.535969	20.74144
Jarque-Bera	1.073759	2.073637	6.678874	3.017045
Probability	0.1200000	0.000000	0.035457	0.000000
Sum	22064754	11348154	13416.97	19.30298
Sum Sq. Dev.	3.64E+13	1.20E+13	4018491	115.3587
Observations	40	40	40	40

The skewness which is a measure of the asymmetry of the distribution of the series around its mean have values greater than 0 which indicates that the series is skewed to the right. This suggests that the distribution has a long right tail. The kurtosis measures the peakedness of flatness of the distribution with an expected value of 3.0. The results in table 1 above shows that the RER and government investment satisfy the condition. However, those of government consumption and PRODUCTIVITY are leptokurtic (greater than 3). The Jarque-Bera test is used to test whether the random variables with unknown means and constant dispersions are normally distributed. The jarque-bera test has the null hypothesis of normally distributed residuals. The probability values indicates an acceptance of the null hypothesis that the errors are normally distributed.

The augmented Dickey Fuller (ADF) and the Phillip Perron (PP) unit root tests were used to test whether the variables are stationary or not and their order of integration. The summary of the ADF and PP unit root tests are shown in table 2 below:

Table 2: Summary of ADF and PP unit root tests

Variables	ADF			PP		
	Level	1st difference	Order of Integration	Level	1st difference	Order of integration
Productivity	3.864607*	-11.06536	1(0)	-6.97216*	-16.78975	1(0)
REER	-2.175824	-4.235913*	1(1)	-2.57216	-4.401870*	1(1)
GINV	-0.826384	-6.790597*	1(1)	-0.508566	-6.369998*	1(1)
GCONS	7.720578*	2.077504	1(0)	12.42145*	-1.672323	1(0)

N.B.: * Indicates statistical significance at the 1% level The results of the unit root tests show that while the REER and GINV were non stationary, PRODUCTIVITY and GCONS were stationary. Following Harris (1995) and Gujarrati (2003), all the variables were carried along for the cointegration test because according to them, both 1(1) and 1(0) variables can be cointegrated. We thus proceed to the next test which is a test of cointegration. The summary of the Johansen cointegration test is shown in table 3 below:

Table 3: Summary of Johansen Cointegration Test Result

Hypothesized No of CE(s)	Eigen Value	Trace Statistic	5% CV	1%CV	Max-Eigen statistic	5% CV	1%CV
None**	0.610914	64.10137	47.21	54.46	34.92636	27.07	32.24
At most 1	0.411748	29.17501	29.68	35.65	19.63222	20.97	25.52
At most 2	0.165128	-9.542795	13.41	20.04	6.677654	14.07	18.63
At most 3	0.07454	2.865141	3.76	6.65	2.86514	3.76	6.65

Both trace and max-eign statistics indicate 1 cointegrating equation at both the 5% and 1% level

The result from the Johansen cointegration test in table 3 above suggests the existence of a long run cointegrating relationship among government consumption, government investment, productivity and the REER in Nigeria. Under this circumstance, favouring a VAR in level or first difference as opposed to Vector Error Correction Model can lead to misspecification because cointegration is established.

The number of cointegrating relationship and the number of lags provided a guide for the specification of VECM. The first step is to identify the number of cointegration relationships that have been suggested in the last section. Table 4 presents the results from the VECM.

Table 4: Summary of Vector Error Correction Result

Vector error correction estimates Date: 01/2/12 time: 09:33 Sample (adjusted): 1970 2010 Included observations: 37 after adjusting endpoints Standard errors in () & t-statistics in []				
CointegratingEq:	cointEq1			
LREER(-1)	1.000000			
LGINV	-0.828758			
	(0.73250)			
	[-1.13142]			
LGCONS(-1)	0.726351			
	(0.63368)			
	[1.14624]			
PRODUCTIVITY	-2.121871			
	(0.39088)			
	[-5.42843]			
C	-3.423876			
Error Correction	D(LREER)	D(LGINV)	D(LGCONS)	D(PRODUCTIVITY)

CointEq1	-0.139607	-0.007866	-0.102223	0.459976
	(0.02444)	(0.03933)	(0.06490)	(0.10177)
	[-5.71282]	[-0.19997]	[-1.57499]	[4.51975
D(REER(-1))	0.205886	-0.057006	0.314901	-0.675435
	(0.19239)	(0.25916)	(0.42762)	(0.67054)
	[1.07013]	[-0.21996]	[0.73638]	[-1.00731]
D(LREER(-2))	-0.066759	0.008332	0.129669	0.041008
	(0.18745)	(0.25250)	(0.41664)	(0.65330
	[-0.35615]	[0.03300]	[0.31123]	[0.06277]
D(LGINV(-1))	-0.058636	-0.052269	0.028703	0.451211
	(0.14480)	(0.19506)	(0.32186)	(0.50468)
	[-0.40493]	[-0.26797]	[0.08918]	[0.89406]
D(LGINV(-2)	-0.055855	-0.110037	0.150028	0.148380
	(0.14099)	(0.18991)	(0.31337)	(0.49137)
	[-0.39617]	[-0.57940]	[0.47875]	[0.30197]
D(LGCONS(-1))	-0.092317	0.156807	-0.419534	0.023679
	(0.08201)	(0.11047)	(0.18228)	(0.28583)
	[-1.12566]	[1.41944]	[-2.30154]	[0.08284]
D(LGONS(2))	0.046628	0.064449	-0.134291	-0.201600
	(0.08678)	(0.11689)	(0.19288)	(0.30243)
	[0.53734]	[0.55137]	[-0.69626]	[-0.66659
(D(PRODUCTIVITY(-1))	-0.100077	0.043718	-0.129730	-0.082509
	(0.05157)	(0.06947)	(0.11462)	(0.17973)
	[-1.94062]	[0.62934]	[-1.13179]	[-0.45907]
D(PRODUCTIVITY(2))				
	-0.027754	0.020305	-0.057117	-0.040316
	(0.03534)	(0.04760)	(0.07855)	(0.12317)
	[-0.78534]	[0.42654]	[-0.72715]	[-0.32733]
C				
	-0.005661	0.170881	0.387890	-0.365514
	(0.06812)	(0.09176)	(0.15141)	(0.23741)
	[-0.08311]	[1.86230]	[2.56190]	[-1.53959]
R-Squared	0.624956	0.106233	0.361972	0.819057
Adj. r-Squared	0.593393	-0.191690	0.149297	0.758743
Sum sq.Resids	2.591077	4.701479	12.80081	31.47331
S.E. Equation	0.309784	0.417287	0.688552	1.079666
F. Statistic	11.87074	0.356578	1.701991	13.57982
Log likelihood	-3.312106	-14.33447	-32.86465	-49.50783
Akaike A/C/	0.718573	1.315377	2.317008	3.216639
Schwarz SC	1.154957	1.750760	2.752391	3.652023
Mean Dependent	-0.044733	0.189643	0.277693	-0.251344
S.D. dependent	0.304737	0.382256	0.746531	2.198109
Determinant Residual Covariance		0.006993		
Log Likelihood		-94.87462		
Log Likelihood (d.f. adjusted)		-118.1906		
Akaike Information Criteria		8.767060		
Schwarz Criteria		10.68275		

A comparison of the coefficients of the error correction terms (Coint Eq1) at the bottom of table 4 for the first vector shows that the LREER has the most significant coefficient and the right sign with a t value of -5.71282. The other variables either have a wrong sign or are not significant. This indicates that the REER

equation constitutes the true cointegrating relationship in the first cointegrating vector. The result thus suggests that about 14 percent of the disequilibrium in the REER is corrected each year. The error correction term for productivity measured by the growth rate of Real Gross Domestic Product, government consumption and government investment are statistically flawed. The error correction for government investment with a value of -0.01 isstatistically significant, while that of productivity 0.46 is wrongly signed and falls outside the acceptance level of -1 < error term < 0. That of government consumption with value of -0.10 falls within the acceptance region, but it is not statistically significant. The result of the VECM thus shows how the REER responds to variation in government investment, government consumption and productivity.

The next stage of this analysis is on the variance decomposition of various variables. The variance decomposition of the variables indicate the percentage of variance as explained by the shock from the variable itself and the shocks from the other variables in the model.

The results of the variance decomposition are shown in tables 5-8 below:

Table 5: Variance Decomposition for GCONS

Period	S.E.	GCONS	GINV	REER	PRODUCTIVITY
1	91401.75	100.0000	0.000000	0.000000	0.000000
2	100655.8	94.84092	0.547241	3.257463	1.354373
3	138890.5	93.12930	0.364911	3.140537	3.365251
4	209880.2	58.34473	34.28373	2.528645	4.842889
5	304003.9	46.12440	46.80287	2.259512	4.813227
6	404208.1	34.74563	57.12309	2.387658	5.743628
7	535313.6	29.03869	62.96274	2.230039	5.768527
8	702186.3	23.02595	69.33604	2020013	5.617999
9	885678.9	2.0.04365	72.50202	1.919947	5.534379
10	1092906	17.37235	75.22525	1.859011	5.543392

Table 6: Variance Decomposition for GINV

Period	S.E.	GCONS	GINV	REER	PRODUCTIVITY
1	2911112.4	0.753798	99..24620	0.000000	0.000000
2	422018.8	1.702683	98.19762	0.099572	1.000127
3	459186.9	1.593918	97.92226	0.472250	0.011569
4	496748.3	1.484823	97.99639	0.502161	0.016629
5	570073.9	1.288804	9828262	0.414390	0.014190
6	619014.0	1.225470	98.34573	0.416352	0.012451
7	652891.3	1.101763	98.41766	0.459170	0.021410
8	693560.3	1.080851	98.44182	0.455714	0.021616
9	739963.9	1.012851	98.52377	0.441826	0.021551
10	775469.6	0.972438	98.55309	0.451661	0.022607

Table 7: Variance Decomposition for REER

Period	S.E.	GCONS	GINV	REER	PRODUCTIVITY
1	82.99009	0.418938	0.326627	99.25423	0.000000
2	138.5861	0.444439	0.119405	99.43030	0.005851
3	173.7214	1.080052	0.099596	98.58227	0.238081
4	196.0991	1.384714	0.187572	98.22732	0.200392
5	216.5612	1.652973	0.161263	98.00702	0.178739
6	239.1067	2.194198	1.358333	96.26110	0.186372
7	263.3880	2.816097	3.699272	93.24691	0.237716
8	287.7562	3.328495	6.063200	90.25697	0.351335
9	314.1670	3.916915	9.418689	86.13450	0.529893
10	344.9682	4.559327	14.29363	80.39402	0.753031

Table 8: Variance Decomposition for Productivity

Period	S.E.	GCONS	GINV	REER	PRODUCTIVITY
1	1.320252	3.804788	0.337901	1.604718	94.25259
2	1.496547	6.853121	3.700279	14.61571	74.83089
3	1.643223	8.914350	4.130312	12.13668	74.81866
4	1.836847	9.272510	11.30787	0.060284	69.55934
5	1.957215	9.611196	12.70665	0.083962	68.51819
6	2.077657	11.56432	11.79388	8.276356	68.36345
7	2.18328	12.11420	11.54346	7.912223	68.43011
8	2.298455	12:71617	11.88527	7.52122	67.93644
9	2.400256	13.40778	11.24064	7.112399	68.23919
10	2.500947	14.40440	10.45250	0.831104	68.31200
Cholesky ordering: GCONS GINV REER Productivity					

A proper assessment of the results reveals that the majority of the variances in the variables have been explained by their own shocks. However, variables such as productivity and government consumption had significant impact on the variance of other variables through the share of their shocks to these variables. The share of the REER to other variables has not been too significant during the period under consideration. The contribution of government consumption shocks to variance in REER increased from 1.4 percent in the 4th period to 4.6 percent in the tenth period, while its contribution to shocks in government investment decreased from 1.7 percent in the 2nd period to about 1 percent in the tenth period. The contribution of government investment to shocks in government consumption was significant at about 75.2 percent in the tenth period, while contribution of government consumption to productivity increased from about 3.8 percent in the 1st period to about 14.4 percent in the last period. Following the result of the variance decomposition of government investment in table 6, other than its own shock, there was no significant contribution of shocks in other variables. The implication of the result is that volatility in government investment and consumption has no major impact on the REER in Nigeria. However, the result indicates that technological productivity is a better target for the government.

The result of the parsimonious ECM result is shown in the appendix. The result indicates that an increase in government investment in the immediate past period depreciated the REER by 0.13 percent. This is an indication that government investment has delivered more productivity gain in the non-tradable sector. The result however indicates that an increase in government consumption by 1 percent appreciated the REER by 0.09 percent. This is an indication that an increase in government consumption increased the relative demand for non-tradable. This result is similar to that by Vahagn and Philip (2008) who used pane data to study the impact of the composition of government spending and RER.

5. Conclusion

The main objective of this study has been to empirically assess the relationship between government expenditure and the REER in Nigeria using a VECM. The theoretical model showed that government consumption typically leads to real appreciation but that sensitivity of the RER to government investment depends on how an increase in the public capital stock differentially affects productivity levels in the traded and non-traded sectors. The descriptive statistic showed that the distribution has a long right tail and that the errors are normally distributed. The result fo the ADF and PP unit root test showed that while two of the variables were stationary, the other two were non-stationary, but were carried along for the cointegration test following Harris (1995) and Gujaratti (2003). The result from the Johansen cointegration test and the VECM suggests a long run relationship among the variables. The result of the variance decomposition showed that no particular portion of the shocks in the REER is attributed to the other variables. However shocks in government expenditures were partly explained by government investment. The result from the parsimonious

ECM model showed that government investment delivered more productivity gains in the non-tradable sector than in the tradable sector. An increase in government consumption increased relative demand for non-tradable sector. The government should thus pay more attention to the expansion of the tradable sector.

References

Coricelli, F. and Jazbec, B., (2004). "Real Exchange Rate Dynamics in Transition economies". *Structural change and Economic Dynamics*. 15: 83 – 100.

Edwards, S., (1989). *Real exchange Rates, Devaluation and Adjustment: Exchange Rate Policy in Developing Countries*. Cambridge: The MIT press.

Elbadawi, I. (1994). "Estimating Long-run Equilibrium Real Exchange Rates".In Williamson, ed., Estimating Equilibrium Exchange Rates, chapter 5. Washington, D.C.: Institute for International Economics. Engle, R. and C. Granger 1987 "Cointegration and error-correction: Representation, estimation, and testing" *Econometrica*, 55: 251 -76.

Engle, R. and C. Granger. (1987). "Cointegrating and Error Correction: Representation, Estimation and Testing". *Econometrica*. 55:251-76

Ghura, D. and Grennes, T.J, (1993). "The Rural Exchange Rate and Macroeconomic Performance in Sub-Saharan Africa". *Journal of Development Economics.* 42: 155-174.

Gujarati, D.N., (2003). *Basic Econometrics* (4e). New York: McGraw-Hill Inc

Harris, R., (1995). *Using cointegration Analysis in Econometric Modeling*. London: Prentice Hall.

Johansen, S; and Juselius, K., (1992). "Testing Structural Hypotehsis in a multivariate Cointegration Analysis of the PPP and the UIP for UK",*Journal of Econometrics*. 53: 211-244.

Joyce, J.P. and Kamas, L., (2003). " Real and Nominal Determinants of Real Exchange Rates in Latin America: Short-run and Long –run Equilibrium". *The Journal of Development Studies,* 39, 6: 155-182.

Macdonald, R., (1998). "What Determines Real Exchange Rates? The Long and Short for it" *Journal of International Financial Markets, Institutions and Money,* 8: 117-1 53

Mkenda, B.K., (2001). "Long-run and Short-run Determinants of the Real Exchange Rate in Zambia". *Goteborg University Working Papers in Economics* No 40. (online). Available:http://www.handels.gu.selepc/archive/00001717/01/gunwe0040.pdf.

Mungule, K.O. (2004). "The Determinants of the Real Exchange rate in Zambia". *African Economic ResearchConsortium, Paper* 146 Nairobi.

Nanna, O.J., Englama, A. and Odoko, F.O. (2004) eds. *Finance, Investment and Growth in Nigeria,* Abuja: Central Bank of Nigeria.

Obadan, M.I. (2006), "Overview of Exchange Rate Management in Nigeria from 1986 to Date", in the "Dynamics of Exchange Rate in Nigeria", *Central Bank of Nigeria Bullion,* Vol. 30, No. 3 pp 17 – 25.

Obstfeld, M. and Rogoff, R. (1998), *Foundations of International Macroeconomics*, Cambridge: MA: MIT Press.

Robert, D.K. and Samuel, J.B (2007) "The Determinants of the Real Exchange Rate in Sierra Leone. JEL Classification Code: FOO,f31,f41.

Vahagn, G. and Philip, R.L. (2008)"The Composition of Government Spending and the Real Exchange Rate, JEL Code E62, F31, F41.

Xiaopu, Z. (2002) "Equilibrium and Misalignment: An Assessment of the RMB Exchange Rate from 1978 to 1999". *Centre for Research on Economic Development and Policy Reforms*, Working Paper No. 127.

Appendix

Summary of Parsimonious ECM Result

Dependent Variable: DLREER
Method: Least Squares
Date: 01/21/12 Time: 11:13
Sample (adjusted): 1970 2101
Included observation: 38 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
DLGINV(-1)	0.128231	0.048192	2.660860	0.0130
DLGONS	-0.085214	0.039559	-2.154078	0.0436
PRODUCTIVITY(-2)	-0.080611	0.026962	-2.98955	0.0053
ECM(-1)	-0.454316	0.167069	-2.719327	0.0105
C	-0.055911	0.059543	-0.939012	0.3545
R-squared	0.8406158	Mean Dependent var		0.045313
Adjusted R-squared	0.8305215	S.D dependent var		0.300612
S.E. of regression	0.308352	Akaike info criterion		0.606932
Sum squared resid	3.137679	Schwarz criterion		0.822404
Log likehood	-6.531713	F-statistics		124.5414
Durbin-Watson stat	2.0468811	Prob(F-statistics)		0.000000

Cohesion Analysis of L2 Writing: The Case of Iranian Undergraduate EFL Learners

Firooz Sadighi, Ph.D.

Department of Foreign Languages, Shiraz Branch,
Islamic Azad University, Shiraz, Iran.
firoozsadighi@yahoo.com

Pooneh Heydari, Ph.D. Candidate

Department of Foreign Languages, Shiraz Branch,
Islamic Azad University, Shiraz, Iran.
corresponding author: Pooneh.Heydari@yahoo.com

Doi:10.5901/mjss.2012.v3n2.557

Abstract: *This study aimed at investigating the most frequent cohesive errors committed by Iranian undergraduate EFL learners at different levels of proficiency as well as the sources of cohesive errors. An overall number of 67 undergraduate students at Shiraz Azad University participated in this study. To have three groups of learners with different proficiency levels, Oxford Placement Test 1B1 (Allan, 1985) was administered. To achieve the objectives of the study, the participants were given a writing task requiring them to write an approximately 200-word narrative composition. Then, the compositions were scored based on the taxonomy developed by Halliday and Hasan (1976). Finally, the data were analyzed through appropriate procedures using quantitative methods. Regarding the frequencies and percentages of errors it was found that low-level learners' most frequent errors were involved in references (20), followed by errors in lexical (14), and conjunctive cohesion (1). Besides, the findings showed that errors in references were the most common (17), followed by errors in lexical (13), and conjunction cohesion (2) in the mid-level learners' narrative compositions and, finally, the high-level learners' most frequent errors were involved in lexical cohesion (17), references (14), conjunction cohesion (3), and substitution (1).This study also allowed for an examination of the sources of cohesive errors. It was found that errors in the use of relative pronouns, conjunctions, along with different forms of repetition appeared because of the incomplete knowledge of the learners—intra-lingual causes. Furthermore, in this study, the errors in the use of personal-, possessive-pronouns, demonstratives and collocations were among the inter-lingual causes of errors.*

Keywords: *Error analysis; Cohesion analysis; Cohesive devices; L2 writing; Cohesive errors; Sources of errors*

Introduction

In recent years, there has been a growing research interest in the analysis of errors adults make while learning a second language. The study and analysis of the errors made by second language learners (i.e. Error Analysis or EA), either in their speech or writing or both has been brought under consideration by many educators, EFL teachers, linguists, and researchers throughout the world. In fact, learners' errors have been the subject of controversy for a long time.

Generally, as Keshavarz (1999, p. 11) stated, "there have been two major approaches to the study of learners' errors, namely Contrastive Analysis and Error Analysis." He further discussed that, "Error Analysis emerged on account of the shortcomings of Contrastive Analysis which was the favored way of describing learners' language in the 1950s and 1960s" (p. 42).

The process involved in CA is the comparison of learners' mother tongue and the target language. Based on the similarities or differences between two languages, predictions were made on errors that learners would be likely or disposed to make as a result (Kim, 2001). As Kim (2001) explained, by early 1970s, CA lost its favor because of the inaccurate or uninformative predictions of learner errors; errors did not occur where predicted, but instead errors showed up where CA had not predicted. More serious criticism was

raised on account of its adopted views from structuralism in linguistics and behaviorism in psychology. Being questioned about the reliability of the CA research, it yielded to Error Analysis in 1970.

Unlike CA which tries to describe differences and similarities of L1 and L2, James (1998 cited in Kim, 2001) stated that, EA attempts to describe learners' interlanguage (i.e. learners' version of the target language) independently and objectively. He believed that the most distinct feature of EA is that the mother tongue is not supposed to be mentioned for comparison. The studies in EA have for the most part dealt with linguistic aspects of learners' errors; not enough attention has been paid to the errors at discourse level and in particular to cohesive devices that are very important in the organization of the texts. Identifying and describing the origin of the learners' errors is now an activity that has received much attention during the last three decades.

According to Halliday and Matthiessen (2004), cohesive devices are formal elements in the text that function to make links between the components of the text. Two broad categories and some subcategories have been identified for cohesive devices. These are grammatical and lexical cohesive devices. The grammatical one includes reference, conjunction, substitute, and ellipsis while the lexical cohesive devices are reiteration and collocation.

Review of the related literature

1. Cohesion Analysis

Of course, studying writing issues involving cohesion deserves much attention. This is because as Ting (2003, p.1) believed, "cohesion as an indispensable text-forming element plays a critical role in composing a text."

The appearance of Halliday and Hasan's *Cohesion in English* (1976) had a major impact on the understanding and teaching of coherence features. Although linguists speak of coherent text as, "having two characteristics such as cohesion (ties between sentences) and register (coherence with a context)" (Choi, 2005), this book focused almost exclusively on cohesion as a text feature. According to Halliday and Hasan (1976), cohesion is realized through cohesive ties which link the presupposing and the presupposed across sentence boundaries. In other words, cohesive ties create intimate intersentential relationships which to a large extent distinguish a text from a sequence of isolated sentences.

Following the publication of Halliday and Hasan (1976), language educators and teachers have become interested in the use of cohesive devices in language students' written compositions. In fact, lots of research has been done in this respect and several researchers have used the outline of cohesion in English presented by various experts particularly the one presented by Halliday and Hasan (1976). Thus, since the present study did so, it is important first to become familiar with the cohesion taxonomy presented by Halliday and Hasan in 1976.

1.1 Halliday and Hasan's (1976) Cohesion Taxonomy

In their classic study of cohesion in English, Halliday and Hasan (1976) defined cohesion as what occurs when the interpretation of some elements in the discourse is dependent on that of another. The one presupposes the other in the sense that it cannot be effectively decoded except by recourse to it. Halliday and Hasan (1976) identified five types of cohesion: reference cohesion, substitution cohesion, ellipsis, lexical cohesion and conjunction cohesion. The first three types fall under the category of grammatical cohesion. Lexical cohesion on the other hand refers to relationships between any lexical item and some previously occurring lexical item in the text quite independently of the grammatical category of the items in question. For example, lexical cohesion can exist between the noun *magistrate* and the verb *judge*. Conjunctive cohesion lies on the borderline between grammatical cohesion and lexical cohesion. Conjunctive cohesion is affected

by cohesion elements that are called conjunctives. This is the Halliday and Hassan's (1976) Cohesion Taxonomy in a glance:

1. Reference cohesion
   - personal reference
     - personal pronouns: *I, me, you, we, us, him, her, they, them, it*
     - personal determiners(the possessives): *my ,mine, your, yours, his, her, hers, their, theirs, its*
     - relative pronouns: *who, which*
   - demonstrative reference
     - determiners: *the, this, there, that, those*
     - demonstrative adverbs: *here, there, then*
   - comparative reference
     - comparative adjectives: *same, identical, equal, other, different, more, better, etc.*
     - comparative adverbs: *similarly, differently, more, less, etc.*
2. Substitution cohesion
   - nominal substitution: *A. Can you give me a glass? B. There is <u>one</u> on the table.*
   - verbal substitution: *Every child likes chocolate and I think my son does too.*
   - clausal substitution: *Latecomers will not be allowed in school after 8.00 a.m. the headmaster says <u>so</u>.*
3. Ellipsis
   - nominal ellipsis: *These are my two dogs. I used to have <u>four</u>.*
   - verbal ellipsis: *Teacher: Have you done the homework? John: Yes, <u>I have</u>.*
   - clausal ellipsis: *Mary: Are you going to buy a new dress for my birthday? Mother: <u>Yes</u>*
4. Conjunctive cohesion: *hence, so, after, and, but, then, etc. e.g. He took a cup of coffee after he woke up.*
5. Lexical cohesion:
   1. repetition of a word or phrase;
   2. synonymy *(e.g. commonly, popularly);*
   3. antonymy *(e.g. high, low);*
   4. hyponymy *(e.g. cigarettes/cigars);*
   5. collocation *(e.g. education, classroom, class, and so on)*

(see Halliday & Hasan, 1976, pp.274-292)

2. Cohesion Studies

As mentioned before, Halliday and Hasan indicated that cohesion is in effect a linguistic property in relation to

textual features. This is while other language researchers have tended to interpret the message of Halliday and Hasan (1976) as follows: given that cohesive devices are important elements for constructing a coherent text, their appearance should cause coherence and therefore contribute to the quality of the text (Ting, 2003). As a result, a number of language researchers adopted Halliday and Hasan's (1976) taxonomy and framework of cohesion to conduct empirical studies examining whether the use of cohesive devices in students' writing correlates with coherence or the overall writing quality. In this section, some of the researches done in this respect will be reviewed briefly.

One of such researchers was Neuuner (1987). He analyzed twenty good essays versus twenty poor essays written by college freshman students. The essays were randomly selected from a pool of 600 essays on the topic "write a letter giving advice to students at school." Two readers from a panel of twelve holistically rated each essay using a four-point scale. Three independent coders conducted analysis on the essays after instruction and practice. Finally, the results revealed that the frequency or percentage of cohesive ties did not distinguish good from weak essays, and good from poor essays did not differ significantly in cohesive distance.

Likewise, using Halliday and Hasan's (1976) model in his study of cohesion and coherence, Khalil (1989) investigated the relationship between cohesion and coherence in 20 compositions in Arab EFL students' college writing. The relationship of cohesion and coherence was tested by the use of multiple correlation statistics. Finally, a weak correlation (r=0.18) was found between the number of cohesive ties and coherence score of the text.

Another researcher was Jonson (1992), who examined cohesion in expository essays written in Malay and in English by native speakers of both languages and in ESL by Malaysian writers. Sample compositions evaluated holistically as good or weak in quality were submitted by Malaysian teachers of composition in Malay and by American teachers of native and non-native speakers of English. The results indicated no differences in the amount of cohesion between good and weak compositions written in Malay by native speakers (20 persons) or in English by native (20 persons) and Malay speakers (20 persons). His empirical study on cohesion in written discourse of native and non-native speakers of English also indicated that judgments of writing quality may depend on overall coherence in content, organization, and style rather than on the quantity of cohesion.

The study of the relationship of cohesion to coherence has continued to dominate the literature of the last two decades. Zhang in 2000 did one such study. He conducted a study to reexamine the same research question by investigating cohesive devices in the writing of Chinese undergraduate EFL students. He asked 107 students of two different universities to participate in his study. The results revealed that no statistically significant relationship exists between the frequency of cohesive ties used and the quality of writing.

Generally, most of the researchers have found that there is no significant relationship between the quantity of cohesive devices used and the quality of writing. However, there are studies in which opposite results were reported. For example, Tonder and Louise (1999), in their study, explored the relationship between densities of lexical cohesion and lexical errors on one hand and the perceived coherence ratings and academic scores of student academic writing on the other. Findings indicated that densities of lexical cohesion generally and derivational ties specifically showed highly significant relations with the coherence ratings.

It has to be pointed out that in any study which intends to examine the relationship between the number of cohesive devices and the quality of writing, counting all the cohesive devices present without taking into account whether or not the cohesive devices are properly used in the context, to some extent makes the study questionable. Furthermore, the inconclusive results reported in the studies reviewed indicate that the relationship between writing quality or textual coherence and cohesive device use has not been concretely established. In the opinion of Castro (2002), such studies yielded conflicting results due to their small sample size, the variability in the L1 subjects involved, and lack of robust statistical analyses to support qualitative descriptions. In short, he continued that cohesion analyses did not consistently reveal differences in cohesive device use in good versus weak essays or between L1 versus L2 writing.

Generally, most of the cohesion analyses have conducted to reexamine the same research question. As it is obvious, despite the significant role of cohesive device in writing, the topic of cohesive errors in composition seems not to have received as much attention as it deserves. There are only a small number of studies exclusively aimed at cohesive errors. Given the fact that cohesive errors have been either neglected or examined incompletely in previous cohesion-related studies, the present study tries to deal with a cohesion analysis of Iranian L2 writing. In doing so, this study attempts to identify the errors using the cohesion taxonomy presented by Halliday and Hasan (1976).

Objectives of the study

The purpose of the present study is to empirically investigate, classify, and analyze the cohesive errors which L2 learners make in their written productions at different levels of proficiency. Moreover, this study tries to investigate whether the identified errors are due to their L2 proficiency level or the L1 interference phenomenon. Hence, the following questions are to be answered through this study:

1. What are the most frequent cohesive errors committed by L2 learners at different levels of proficiency?
2. Are there any differences in the L2 learners' cohesive errors which can be attributed to their L2 proficiency level?
3. Are there any differences in the L2 learners' cohesive errors which can be attributed to L1 interference phenomenon?

Method

Participants

67 male and female EFL undergraduate students at Shiraz Azad University participated in this study—42 female and 25 male. The participants ranged in age from 20 to 26. All of the students speak Persian as their native language and learn English as a foreign language. The type of sampling involved non-random procedures for selecting the members of the sample. In other words, the selection procedure was a non-probability one. The specific selection procedure employed was that of convenient sampling.

Instruments

To have three groups of learners with different proficiency levels, Oxford Placement Test 1B1 which is a standard test including 50 multiple choice items identifying and assessing the learners' level of English proficiency (Allan, 1985) was administered. The validity of the Oxford Placement Test 1B1 is taken for granted and with regard to the reliability of the test, the Kurder-Richardson formula 21 was used and the results showed the reliability of 0.86.

In order to conduct the study, the participants were asked to write a composition. The composition is beneficial in such cases because it will bring naturally occurring data for the study. Instead of administering multiple-choice exams such as tests of grammar or vocabulary that draw the students' most attention and consciousness towards the grammatical and lexical points, the composition test draws the attention of the students towards the topic. In this way, the participants are unaware of grammatical and lexical issues and focus on the subject they want to develop.

Accordingly, the participants were given a writing task requiring them to choose one of the three presented topics and write an approximately 200-word narrative composition. The three narrative topics were as follows:

1. A time and a place in the past
2. A melodic memory

3. A one-day visit to your country

The mentioned topics were selected from the e-book of *Answers to All TOEFL Essay Questions* by ToeflEssays.com, an e-book containing 450 model essays which offers an intensive preparation for the TWE test.

Data collection procedures

The scoring of Oxford Placement Test was based on the number of items answered correctly by the students. Each correct answer received one point and the total score of the test was 50. The participants were first assigned to three groups of low, mid, and high based on the results of the oxford placement test—22 low-level learners, 27 mid-level learners, and 18 high-level learners. The criterion for this division was the standard deviation. Accordingly, participants with one standard deviation below the mean were assigned the low group and the ones with scores falling one standard deviation above the mean were assigned the high group. The remaining ones in the middle were assigned the mid group.

After dividing the participants into three different levels of proficiency, they were given two optional narrative topics on which they were required to write a composition in about sixty minutes. In selecting the topics from the e-book of *Answers to All TOEFL Essay Questions*, two factors were of concern: (1) learners' familiarity with the topics and (2) the topics being interesting and easy to write. The reason for choosing narrative topics was that it was found to be the easiest among other modes of discourse for the learners to write (Nemati, 1999). The data for this study were collected from compositions written by the students during one session. Then, they were scored based on the taxonomy developed by Halliday and Hasan (1976). Consequently, all five cohesive devices consisting of reference (with differentiation made between pronominal, demonstrative, and comparative), substitution (discriminating between nominal, verbal, clausal), ellipsis (discriminating between nominal, verbal, clausal), conjunction and lexical cohesion were scored.

Data analysis procedures

The data were analyzed through appropriate procedures using both qualitative and quantitative methods.

For the first research question of the study investigating the most frequent cohesive errors committed by EFL undergraduate learners, the data were analyzed through descriptive statistics using frequencies and percentages.

For the next two research questions of the study investigating if there are any differences in the participants' cohesive errors which can be attributed to either their L2 proficiency level or L1 interference phenomenon, the data were analyzed using qualitative methods.

Results

1. the most frequent cohesive errors committed by L2 learners at different levels of proficiency

As mentioned before, in order to analyze cohesive errors, the participants were asked to write narrative compositions. The five major categories explicated by Halliday and Hasan (1976) were used to systematically present a framework for the analysis of cohesion in this study. Based on these groupings, common error types which led to their deviation from standard English usage were established. The following discussion will present the major error types, those which were frequent at different levels of proficiency, separately.

1.1. Low-level learners' most frequent cohesive errors

Investigating 22 compositions written by low-level learners identified the 35 cohesive errors in which the use of references were the most frequent ones (20), followed by errors in lexical (14), and conjunctive cohesion (1).

The results presented in Table 1 indicate the relative degree to which the items from each category and subcategory of cohesion were used incorrectly in creating overall cohesion together with a detailed discussion of the error analysis.

Table 1: The low-level learners' frequencies and percentages of errors in cohesion categories and subcategories

Cohesion Categories	Number of errors	Percentages of errors	Cohesion Subcategories	Number of errors	Percentages of errors
Reference	20	57.14%	Personal Demonstrative Comparative	19 1 0	95% 5% 0%
Substitution	0	0%	Nominal Verbal Clausal	0 0 0	0% 0% 0%
Ellipsis	0	0%	Nominal Verbal Clausal	0 0 0	0% 0% 0%
Conjunction	1	2.58%		1	
lexical	14	40%	Repetition Synonymy Antonymy Hyponymy Collocation	4 0 0 0 10	28.57% 0% 0% 0% 71.42%

As is clear from the table, the majority of the errors are referential cohesive devices and the second most frequent errors are those of the lexical type. Moreover, it is clear that the majority of the referential errors are personal and the majority of the second most frequent errors, lexical type, are involved in collocation and repetition, respectively. Examples from each of the error types are also presented in Table 2. In each of the examples, the devices which were used incorrectly are underlined.

Table 2: Low-level learners' cohesive errors

Categories	Subcategories	Low-level learners' cohesion errors
1. Reference	Personal (Pronoun)	1. She wanted to come home but <u>he</u> couldn't. (she) 2. My mother had to go because <u>he</u> was a teacher. (she) 3. My brother went to school, too. <u>She</u> was happy. (he) 4. Some of them were happy but some were not. Among <u>those</u> was a crying girl. (them) 5. The old woman went to the hospital. <u>He</u> was sick. (she) 6. They invited their friends to the party but <u>he</u> didn't came. (they) 7. She heard the news from TV. …she talked about <u>them</u> with his neighbor. (it)
	Personal (Determiners)	1. She talked about them with <u>his</u> neighbor. (her)

		2. My sister passed the course but <u>his</u> score was low. (her) 3. When they were in Iran, they saw many things that they were not in <u>his</u> country. (their) 4. We enjoyed from the movie and we talked about <u>their</u> events with ourselves. (its) 5. When we want to migrate to another country, we might try to learn the culture and match <u>themselves</u> with them. (ourselves) 6. …keep in touch by his own family and friends at <u>their</u> original country. (his) 7. <u>Her</u> wife saw the man which was thief. (his)
	Relative Pronoun	1. Her wife saw that man <u>which</u> was thief. (who) 2. Education is the important criteria <u>who</u> can change the personality of person. (that) 3. At first, I was introduced to a nice woman <u>which</u> became my teacher. (who) 4. …or call with the people <u>which</u> live there. (who) 5. They didn't have a lot of time to think about the country <u>that</u> they born. (where)
	Demonstrative	1. To have some good friends is the best way to prevent <u>this</u> bad problems. (these)
	Comparative	
2. Substitution	Nominal Verbal Clausal	
3.Ellipsis	Nominal Verbal Clausal	
4. Conjunction		1. We can see <u>the</u> our personality has more effect in our life. (that)
5.Lexical	Repetition	1. After that I <u>returned back</u> to my home. (returned home) 2. The most interesting day of my life <u>it</u> is the first day of my school. (repetition of a noun by its pronoun) 3. The characteristics that we born with <u>it</u> is very important. 4. When they were in Iran, they saw <u>many things</u> that <u>they</u> were not in his country.
	Synonymy	

Antonymy	
Hyponymy	
Collocation	1. Foreign people have a different way to life. (way of life) 2. For reduce the problems we can use of foreign web-site or… (make use of) 3. I keep in touch by my family. (keep in touch with sb) 4,5. …keep in touch by his own family and friends at their original country. (keep in touch with sb, home country) 6. I agree with this issue. (agree on sth) 7. We can see the our personality has more effect in our life. (effect on sb/sth) 8. We enjoyed from the road. (enjoy sth) 9. We enjoyed from the movie. (enjoy sth) 10. I run to home from my school. (run home)

1.2. Mid-level learners' most frequent cohesive errors

Besides, a total of 32 cohesive errors in the narrative compositions of 27 mid-level learners were identified. The findings showed that errors in references were the most common (17), followed by errors in lexical (13), and conjunction cohesion (2). The frequencies and percentages of errors from each category and subcategory are presented in Table 3.

Table 3: The mid-level learners' frequencies and percentages of errors in cohesion categories and subcategories

Cohesion Categories	Number of errors	Percentages of errors	Cohesion Subcategories	Number of errors	Percentages of errors
Reference	17	53.21%	Personal Demonstrative Comparative	15 2 0	88.23% 11.76% 0%
Substitution	0	0%	Nominal Verbal Clausal	0 0 0	0% 0% 0%
Ellipsis	0	0%	Nominal Verbal Clausal	0 0 0	0% 0% 0%
Conjunction	2	6.25%		2	
lexical	13	40.62%	Repetition Synonymy Antonymy Hyponymy Collocation	4 0 0 0 9	30.76% 0% 0% 0% 69.23%

As this table show, the majority of the errors are in references and the second most frequent errors are in lexical type. Table 3 reveals that the majority of the referential errors are personal and the majority of the second most frequent errors, lexical errors, are involved in collocation and repetition, respectively. Examples from each of the error types are shown in Table 4. In each of the examples, the devices which were used incorrectly are underlined.

Table 4: Mid-level learners' cohesive errors

Categories	Subcategories	Mid-level learners' cohesion errors
1. Reference	Personal (Pronoun)	1. I think if a single woman live abroad, <u>they</u> will face many problems. (she)
		2. My father gave me bunch of flower…<u>she</u> told me…(he)
		3. Mina had a valuable experience…<u>he</u> remember it after that.(she)
		4. My father couldn't start the car. <u>She</u> decided to…(he)
		5. I saw a girl that were kind in the first day of school. <u>He</u> became my friend after that. (she)
		6. My friend told teacher that <u>she</u> was wrong. (confusing pronoun)
	Personal (Determiners)	1. Finally he could found a way to improve his life of <u>hisself</u> and children. (himself)
		2. If a student is worried all the time <u>their</u> scores will be not good. (his/her)
		3. Teacher told us that we should study hard to good scores in <u>my</u> exams. (our)
		4. my grandmother was sick. I bought <u>his</u> drugs. (her)
	Relative Pronoun	1. The students <u>which</u> know English have a better chance for work. (who)
		2. The people <u>that</u> they saw this movie…(who)
		3. This problems is worse for women especially <u>that they</u> live in the countries like Iran. (who)
		4. the topic <u>in which</u> the teacher talked was…(that)
		5. I saw a girl <u>that</u> were kind in the first day of school. (who)
	Demonstrative	1. This <u>problems</u> is worse for women especially that they live in the countries like Iran. (problem)
		2. The husband's idea differed from wife's idea about <u>this</u> issues.(these)