

Modelling South African House Prices: A bubbly story or not?

Lizelle Janse van Rensburg¹

This is a work in progress. Please do not cite.

Abstract: There has been much speculation in recent times concerning the rapid rise in South African house prices. In 2004 real year-on-year residential property price growth was 30.3% which represented the largest increase in the national average residential property price growth since 1981. Consequently many commentators suggested that there is a speculative bubble in the South African housing market. To answer this question the paper models residential property prices in South Africa using fundamental factors consistent with economic theory. If a bubble is defined as a large unexplained change in property prices, and if these fundamental factors can explain property prices, then no bubble exists. In estimating the model of real house prices, the paper considers using a Johansen Vector error-correction framework (1988, 1990, 1992). For the middle segment of the housing market, findings meet some of the a priori theoretical expectations when considering the relationship between real house prices and real disposable income, the mortgage rate, the inflation rate and expectations as indicated by the business confidence index.

JEL: D40; L85

1. Introduction

House prices have more than tripled since 1999. In 2005 it was reported that real house prices grew uninterruptedly in South Africa for the fifth consecutive year. Nominal year-on-year residential property price growth was 32.2% or in real terms 30.3% in 2004, which was the largest increase in the national average residential property price growth since 1981 (ABSA: 2005:2). The sharp rise in house prices recently may prompt the question of whether or not there is a bubble in the housing market, i.e. whether or not house prices are far higher than a fundamental value determined by interest rates, income and other explanatory factors

¹ Department of Economics, University of the Free State

for house prices. According to Jacobson (2005:29) a house price bubble can arise if many individuals purchase a dwelling during the current period based on an expected rise in house prices during the following period. Expectations are assumed not to be based on economical fundamentals, sharp adjustment of prices follow when either or both expectations and fundamentals change.

A reversal of house price misalignment would be likely to have strong repercussions on the economy. For most individuals their residential property is the primary component of their wealth (Riddle, 1999:273). A sudden decrease in property value may cause banks to experience that the value of the collateral falls below the value of the loan and that households increasingly have difficulty repaying debt. The downturns in the property cycles not only affects consumption decisions but can also lead to bankruptcies in the financial and corporate sectors, along with macro-economic recession and unemployment (Riddle, 1999:273; Weeken, 2004:1, Case 2000:120, 148; Jacobsen, 2005:29). While there is little doubt that the rates of increase observed in the residential property sector are unsustainable, there is uncertainty as to the sustainability of the level of house prices, or the likelihood of price to fall. If, however, the level of house prices can be explained by fundamental factors, house prices are not experiencing a market bubble.

The purpose of this paper is to model the long run cointegrated relationship between house prices in South Africa and the fundamental factors consistent with economic theory. Specifically the analysis utilizes Johansen Vector error correction framework (1988, 1990, 1992) which remedy in part some of the modelling problems when estimating residential price movements. The argument behind this modelling procedure is that if a long-run relationship between house prices and economic fundamentals exists and the error-correction model can explain deviations from the long-term equilibrium, then no bubble exists.

2. Theoretical Framework

The main theoretical factor that “drives” house prices is the level of disposable income (Jacobson, 2005:37; Foley: 2004, 16 (Leishman:2003:137; Stevenson, 2005:16). Other variables that are commonly included as fundamental factors in modelling approaches are

- the real after tax borrowing rate,
- a demographic indicator such as the household formation rate, an indicator of growth in demand, or the number of households in certain age groups, often the 20-29 age group as such households are particularly economically active,
- a housing stock indicator,
- employment or unemployment indicators,
- various forms of inflation indicators, and
- expectations proxies (Jacobson. 2005:33 ; Stevenson, 2005:10; Foley, 2004:20; Muellbaur and Murphy, 1997:1716)

Porterba (1984: 729-752) and Muellbauer and Murphy (1997:1701-1727)² have put forward models of the USA and UK housing market respectively that equate housing demand to housing supply. In these models house price volatility reflects the movement towards a new equilibrium price following a demand or a supply shock. Deviations from the market-clearing equilibrium price are accentuated by large transaction costs and inertia in the response of housing supply to shocks. In these models house prices are driven only by fundamentals and any deviations are regarded as both rational and efficient. Therefore prices will tend to mean-revert and observed house price movements will reflect a combination of shocks and adjustment mechanisms. The adjustment mechanisms imply that house prices will display positive autocorrelation. In addition these models are usually estimated in a static manner over the entire sample period and fail to fully capture short-run deviations from fundamental value (Stevenson, 2005:19).

Modelling house prices and economic variables is often associated with problems of endogeneity of regressors leading to simultaneity among the variables in the system. In addition macroeconomic variables are aggregated and exhibit non-stationarity. The conventional wisdom was to difference all non-stationary variables used in a regression

² A similar analysis performed by Foley (2004:1-38) on Irish house prices and fundamental determinants reported positive serial correlation even after various adjustments to the model to remedy the problem.

analysis. However, differencing variables to get them stationary eradicates the long-run relationship between the variables (Masih and Masih 1996:321). However, the application of cointegration techniques and developments in error-correction modelling serves, in part, as a solution to these modelling problems.

Jacobsen (2005: 29-40) modelled the Norwegian housing market in an error correction framework with quarterly data from 1990 to the first quarter of 2004. In analysing the factors that caused house prices to rise, the paper finds that interest rates, housing construction, unemployment and household income are the most important explanatory factors for house prices in the Norwegian market. The analysis suggests no evidence of overvaluation in Norwegian house prices compared with the fundamental value of the variables.

Abelson and Joyeux (2005:1-29) estimated a long-run equilibrium model and a short-run asymmetric error correction model of Australian house prices. The authors show that house prices are determined significantly by real disposable income, the consumer price index, unemployment, real mortgage rates, equity prices and the supply of housing. The short run asymmetric error correction model reveals that the housing market adjusts to equilibrium in four quarters when real house prices are rising at more than 2 per cent per annum and 6 quarters when real house prices are flat or falling.

The model used in this paper assumes that the fundamental value of large-, medium-, and small houses in the middle segment category is related to household income and the interest rate, thereby allowing the following empirical specification.

$$P_t = \mu + \phi_t RDY_t + \gamma_t RMOR_t + \varepsilon_t \quad (1)$$

where P_t denotes real house prices, RDY real disposable income, and $RMOR$ denoting the real mortgage rate.

The short-run dynamics are:

$$\Delta P_t = \rho \sum_{j=1}^{j=k} \theta_j \Delta P_{t-j} + \sum_{j=0}^{j=1} \eta_j \text{BER}_{t-j} + \sum_{j=0}^{j=1} \lambda_j \Delta \text{CPI}_{t-j} + \sum_{i=1}^r \xi_i (P_t - (\mu + \phi_t \text{RDY}_t + \gamma_t \text{RMOR}_t)) + \varepsilon_t \quad (2)$$

where BER denotes the business confidence index, and CPI the consumer price index for metropolitan and other urban areas.

3. Econometric methodology and data

(i) Data

The house price time series are obtained from the ABSA residential property market database. According to the ABSA residential property market database the South African residential property market is categorised into three major segments: luxury houses, middle-segment houses (subdivided into large -, medium - and small houses) and affordable houses. The luxury segment refer to houses valued between R2.6 million and R9.5 million); Middle segment houses refer to houses that are valued up to R2.6 million and are further divided into sub segments according to the size of the property. Small houses extend over an area of 80m² to 140m², medium houses 141m² to 220m² and large houses 221m² to 400m². Affordable housing refers to houses of 40m² to 79m² and priced at R193 000 or less.³

Table 1 – Categories of houses in ABSA residential property market data base

Categories	Price	Area (m ²)
Luxury housing	R2.6 million <P _H > R9.5 million	
Middle Segment housing	R226 000 <P _H > R2.6 million	
<i>Small</i>		80 – 140
<i>Medium</i>		141 – 220
<i>Large</i>		221 – 400

³ Due to the increase in house prices, the cut-off prices for the classification of houses were adjusted in 2006. Prior 2006 the price bracket for luxury housing was between R2.2 million and R8.2 million. Middle segment housing were priced up to R2.2 million and affordable housing were priced at up R193 000.

Affordable housing	$P_H < 226\,000$	40 – 79
--------------------	------------------	---------

To model residential property prices in South Africa the paper employs house prices that are categorised in the middle segment of the residential property market separately as large ($P_{large,t}$), medium ($P_{medium,t}$), and small ($P_{small,t}$) houses. The data consists of annual data for the period 1980 to 2006. Prior to 1980 financial market regulations inhibited the smooth functioning of financial markets but from around 1980 onwards the financial markets were liberalised and experienced more freedom.

Based on the literature reviewed above and the data available this paper uses the predominant rate of new mortgage loans extended by banking sector for dwellings and disposable income per capita of households to explain long-run house prices. For the short run dynamics the model uses the Consumer Price Index for metropolitan and other urban areas and the Business Confidence Index.

(ii) Econometric methodology

To estimate the long-run relationship this paper uses a Johansen Vector error-correction framework (1988, 1990, 1992). The analysis examines the long run relationship between house prices (specifically large- medium- and small houses in the middle segment of the residential property market), real disposable income and the real mortgage rate.

4. Estimation results and evaluation

(i) Univariate Time series characteristics of the data

To test for the presence of unit roots the analysis employs the augmented Dickey-Fuller test (ADF) (Dickey and Fuller, 1979) and the Phillips-Perron test (PP) (Phillips-Perron, 1988). For the ADF test, the lag length was chosen using the Modified Schwartz information criterion and the PP test used the Newey-West Bandwidth for lag length determination. It should be noted that these unit root testing procedures can suffer from lack of power as they

tend to accept the null of a unit root too frequently against a stationary alternative (Masih and Masih, 1996: 321). To test for mean stationarity the study performed the test procedure proposed by Kwiatkowski et al. (1992) (KPSS) that tests the null hypothesis of mean stationarity against the alternative. Table 2 shows the results for the ADF, PP and KPSS tests of the variables that were considered in the modelling process.

Given the weakness of the ADF test relative to the PP test, and the weakness of the PP test relative to the KPSS test, this paper will give precedence to the KPSS test over the PP test and the PP test over the ADF test. This approach is similar to that of (Masih and Masih, 1996: 321).

Table 2 – ADF and PP tests for unit roots and KPSS tests for mean stationarity.

Variable	Augmented Dickey-Fuller		Phillips-Perron		KPSS	
			Z(ta)	Z(ta)	h(m)	h(m)
	I(0)	I(1)	I(0)	I(1)	I(0)	I(1)
LNP _{large}	0.72	-2.04**	0.87	2.10**	0.18**	0.17
LNP _{medium}	0.72	-2.14**	0.78	-2.25**	0.18**	0.16
LNP _{small}	0.70	-2.21**	0.74	-2.21**	0.18**	0.15
LNDY _{per capita}	-1.58	-1.14	1.20	-3.61***	0.19**	0.16
LNCPPI	-0.31	-2.60	4.12	-1.42	0.22**	0.13
BER _{business}	0.13	-3.78***	-0.31	-3.76***	0.46	
RMOR _{rate}	-0.28	-4.77***	-0.99	-4.76***	0.10**	0.04

*, **, *** indicates statistical significance at the 10%, 5% and 1% levels respectively

For all variables, except for the consumer price index, the observed value of the PP test is quite small which suggest that the variables contain a unit root at conventional levels of statistical significance. On the other hand the observed value of the KPSS test statistic $\eta(\mu)$ is large at a 5% level of statistical significance and therefore suggests that the variables have a unit root and are non-stationary in levels. The application of the PP test on the first difference of these variables indicates that all variables, except for the CPI, are stationary in their first differences, while being non-stationary in their level form. The PP test indicates that the CPI is non-stationary in first difference form while the KPSS test suggests that the first difference of the CPI is stationary. However, on apriori grounds the paper imposes an assumption of stationarity on the inflation rate. Although the observed value of the ADF and PP test suggest that $Ber_{business}$ contain a unit root and is an I(1) variable the paper gives precedence to the KPSS test and will therefore assume that $Ber_{business}$ is stationary in levels.

4. Estimation results

The paper employs the Johansen and Juselius (JJ) procedure of testing for the presence of multiple cointegrating vectors (LNP_{large} $LNDY_{per\ capita}$ $RMOR_{rate}$), (LNP_{medium} $LNDY_{per\ capita}$ $RMOR_{rate}$), and (LNP_{small} $LNDY_{per\ capita}$ $RMOR_{rate}$). Results of the JJ maximal eigenvalue and trace tests are presented in Table 3 for the three vector specifications.

Table 3 – Cointegration based on the Maximal Eigenvalue and Trace test

<i>Vector</i> (LNP_{large} $LNDY_{per\ capita}$ $RMOR_{rate}$)			<i>95% Critical</i>		<i>95% Critical</i>
<i>Null</i>	<i>Alternative</i>	<i>Eigenvalue Statistic</i>	<i>Value</i>	<i>Trace Statistic</i>	<i>Value</i>
$r = 0$	$r \geq 1$	43.26*	22.3	57.39*	35.19

$r \leq 1$	$r \geq 2$	13.03	15.89	14.13	20.26
$r \leq 2$	$r \geq 3$	1.10	9.16	1.10	9.16

Vector (LNP_{medium} LNDY_{per capita} RMOR_{rate})

<i>Null</i>	<i>Alternative</i>	<i>Eigenvalue Statistic</i>	<i>95% Critical Value</i>	<i>Trace Statistic</i>	<i>95% Critical Value</i>
$r = 0$	$r \geq 1$	47.34*	47.34	56.48*	35.19
$r \leq 1$	$r \geq 2$	8.22	8.22	9.14	20.26
$r \leq 2$	$r \geq 3$	0.92	0.92	0.92	9.16

Vector (LNP_{small} LNDY_{per capita} RMOR_{rate})

<i>Null</i>	<i>Alternative</i>	<i>Eigenvalue Statistic</i>	<i>95% Critical Value</i>	<i>Trace Statistic</i>	<i>95% Critical Value</i>
$r = 0$	$r \geq 1$	35.19*	35.19	47.59*	35.19
$r \leq 1$	$r \geq 2$	20.26	20.26	6.61	20.26
$r \leq 2$	$r \geq 3$	9.16	9.16	0.92	9.16

* denotes statistical significance.

The JJ maximal eigenvalue and trace tests indicate the presence of at most a single cointegrating vector among the $(LNP_{large} \text{ LNDY}_{per capita} \text{ RMOR}_{rate})$, $(LNP_{medium} \text{ LNDY}_{per capita} \text{ RMOR}_{rate})$ and $(LNP_{small} \text{ LNDY}_{per capita} \text{ RMOR}_{rate})$ variables.

Estimates of the long run parameters are:

$$LNP_{large} = 1.87LNDY_{per capita} - 0.05RMOR_{rate} \quad (3)$$

[2.74*] [6.52*]

$$LNP_{medium} = 1.89LNDY_{per capita} - 0.07RMOR_{rate} \quad (4)$$

[2.43*] [7.22*]

$$LNP_{small} = 1.19LNDY_{per capita} - 0.10RMOR_{rate} \quad (5)$$

[0.97] [7.30*]

The results are consistent with the theoretical priors and all the parameters are statistically significant, except for the household disposable income parameter of small middle segment houses (Equation (5)), so that only the mortgage rate parameter is statistical significant.

To analyse the relative importance of each random innovation that affects the variables in the VAR, the paper performed variance decomposition of the VAR variables with the real mortgage rate as first shock in the Cholesky ordering. According to the variance decomposition (at around 4 years), real disposable income explains only 2.14% of the variance of real large house prices, while the real mortgage rate explains 62.44%. Large house prices explain 35.42% of the variance in itself. The variance decomposition for medium houses suggest that disposable income explain only 5.12% and the mortgage rate explain 75.78% of the variance in medium house prices. 19.04% of the variance in medium house prices can be explained by the variance in itself. The variance in small house prices are largely explained by the mortgage rate (57.89%) and the variance of small house prices (39.65%) itself. Disposable income is responsible for only 2.46% of the variance in small house prices.

The dynamics of the error correction models for LNP_{large} , LNP_{medium} and LNP_{small} follow in Table 4.

Table 4 – Error Correction representation.

	Dependent variable = ΔLNP_{large}	Dependent variable = ΔLNP_{medium}	Dependent variable = ΔLNP_{small}
$\Delta LNP_{large}(-1)$	0.68 [3.49]	-	-
$\Delta LNP_{large}(-2)$	-0.22 [0.88]	-	-
$\Delta LNP_{large}(-3)$	0.38 [2.15]	-	-
$\Delta LNP_{medium}(-1)$	-	0.67 [3.74]	-
$\Delta LNP_{medium}(-2)$	-	-0.18 [0.75]	-
$\Delta LNP_{medium}(-3)$	-	0.41 [2.25]	-
$\Delta LNP_{small}(-1)$	-	-	0.60 [3.01]
$\Delta LNP_{small}(-2)$	-	-	-0.05 [0.21]
$\Delta LNP_{small}(-3)$	-	-	0.17 [0.84]
$\Delta LNDY_{per\ capita}(-1)$	-0.06 [0.15]	0.78 [1.78]	0.66 [1.24]
$\Delta LNDY_{per\ capita}(-2)$	-0.21 [0.61]	-0.01 [-0.02]	0.60 [1.54]
$\Delta LNDY_{per\ capita}(-3)$	0.56 [1.56]	0.66 [1.62]	0.40 [0.96]
$\Delta RMOR_{rate}(-1)$	0.00 [0.54]	0.00 [0.33]	0.00 [0.05]

$\Delta\text{RMOR}_{rate(-2)}$	0.01 [1.95]	0.01 [2.14]	0.01 [2.22]
$\Delta\text{RMOR}_{rate(-3)}$	0.00 [0.41]	0.00 [0.06]	0.00 [0.77]
ΔLNCPPI	-1.23 [4.38]	-1.58 [4.88]	-2.14 [4.91]
$\text{BER}_{business}$	0.00 [2.30]	0.00 [0.25]	0.00 [0.44]
ECM	-0.27 [3.09]	-0.31 [4.37]	-0.25 [4.14]
Adjusted R ²	0.84	0.84	0.79
F-stat	13.12	13.04	9.72
SE of Regression	0.04	0.04	0.05
Mean of dependent variable	0.03	0.03	0.02
SD of Dependent Variable	0.10	0.11	0.11
Residual Sum of Squares	0.03	0.03	0.04
Equation Log-Likelihood	55.89	54.33	51.11
Serial Correlation $\sim \chi^2(3)$	14.49	7.06	9.78
Normality $\sim \chi^2(2)$	5.59 (0.06)	4.67 (0.1)	6.20 (0.05)
Heteroskedasticity $\sim \chi^2(1)$	143.14 (0.50)	142.47 (0.52)	142.88 (0.51)

27 Observations included over the period 1980 to 2006. Square parentheses denote t-values and round parentheses denote probability. ECM denotes error correction term obtained from the long run relationship.

The error correction term obtained from the long run relationship for LNP_{large} , LNP_{medium} , and LNP_{small} is statistically significant in all three models. The short-run parameter of the inflation rate meet the a priori expectations and show that a percentage increase in the inflation rate will affect the price change in all three sub-segments (large-, medium- and small house prices) negatively with between 1.23% and 2.14%. Although the short run parameter of the business confidence index is statistical significant in the case of large middle segment house prices, it is of no economical value due to the small size of the parameter. The business confidence parameter is not statistical significant in the short-run dynamics of medium- and small house prices.

5. Conclusion

The paper attempts to model the middle segment of the residential housing market with available data series consistent with economic theory. The above analysis indicates that developments in house prices in recent years can be largely attributed to changes in

fundamentals. The analysis so far does not provide evidence that house prices are overvalued in terms with economical determinants. Thus, there is no evidence that a market bubble exist. The paper provides evidence that the long-run price level of large and medium houses in the middle segment of the residential property market can be explained by the theoretical determinants that is real disposable income and the real mortgage rate. Short-run dynamics are explained mainly by the inflation rate, while business confidence seems only to affect large house prices and not medium house prices. In modelling small house prices only the mortgage rate is statistical significant in the long-run and only the inflation rate statistical significant at conventional levels in the short-run.

6. Bibliography

Abelson, P., R. Joyeux, et al. (2005). "House Prices in Australia: 1970 to 2003 Facts and

- Explanations." **4**: 1-29.
- Case, K., E. Glaeser, et al. (2000). "Real Estate and the Macroeconomy." Brookings Papers on Economic Activity **2000**(2): 119-162.
- Dickey, D. and W. Fuller (1979). "Distribution of the estimators for autoregressive time-series with a unit root." Journal of the American Statistical Association **74**: 427-31.
- Foley, P. P. (2004). "Are Irish house prices determined by fundamentals?" Working paper **1**(2004).
- Jacobsen, D. H. (2005). "What drives house prices?" Economic Bulletin **Q1**: 29-39.
- Johansen, S. (1985). "Understanding Inflation in the 1980's." Brooking Papers on Economic Activity **85**: 263-99.
- Johansen, S. (1988). "Statistical analysis of cointegrating vectors." Journal of Economic Dynamics and Control **12**: 551-70.
- Johansen, S. and K. Juselius (1992). "Testing structural hypotheses in a multivariate cointegrating analysis of the PPP and the UIP for UK." Journal of Econometrics **53**: 211-44.
- Kwiatkowski, D., P. Phillips, et al. (1992). "Testing the null hypothesis of stationarity against the alternative of a unit root: how sure are we that economic time series have a unit root?" Journal of Econometrics **54**: 157-178.
- Leishman, C. (2003). Real Estate Market Research and Analysis. New York, Palgrave Macmillan.
- Masih, R. and A. M. Masih (1996). "Stock-Watson dynamic OLS (DOLS) and error-correction modelling approaches to estimating long- and short-run elasticities in a demand function: new evidence and methodological implications from an application to the demand for coal in mainland China." Energy Economics **18**: 315-334.
- Muelbauer, J. and A. Murphy (1997). "Booms and Busts in the UK Housing Market." The

Economic Journal **107**(November): 1701-1727.

Phillips, P. and P. Perron (1988). "Testing for a unit root in time series regression."
Biometrika Theory **7**: 1-21.

Poterba, J. (1984). "Tax Subsidies to Owner-Occupied Housing: An Asset-Market
Approach." The Quarterly Journal of Economics **99**(4): 729-752.

Riddle, M. (1999). "Fundamentals, Feedback Trading, and Housing Market Speculation:
Evidence from California." Journal of Housing Economics **8**: 272-284.

Stevenson, S. (2005). Modelling Housing Market Fundamentals: Empirical Evidence of
Extreme Market Conditions. Housing Fluctuations in Europe: Causes and Policy
Options Conference, Madrid.

Weeken, O. (2004). "Asset Pricing and the housing market." Bank of England Quarterly
Bulletin **Spring 2004**: Available at SSRN: <http://ssrn.com/abstract=701272>.