

Forecasting US Real House Price Returns over 1831-2013: Evidence from Copula Models

Rangan Gupta* and Anandamayee Majumdar**

Abstract

Given the existence of non-normality and nonlinearity in the data generating process of real house price returns over the period of 1831-2013, this paper compares the ability of various univariate copula models, relative to standard benchmarks (naive and autoregressive models) in forecasting real US house price over the annual out-of-sample period of 1859-2013, based on an in-sample of 1831-1873. Overall, our results provide overwhelming evidence in favor of the copula models (Normal, Student's t , Clayton, Frank, Gumbel, Joe and Ali-Mikhail-Huq) relative to linear benchmarks, and especially for the Student's t copula, which outperforms all other models both in terms of in-sample and out-of-sample predictability results. Our results highlight the importance of accounting for non-normality and nonlinearity in the data generating process of real house price returns for the US economy for nearly two centuries of data.

Keywords: House Price; Copula Models; Forecasting
JEL Codes: C22, C53, R3

1. Introduction

The recent global financial crisis has resulted in a keen interest in both academic and policy circles on the role of asset prices, and in particular, housing prices on economic activity. As Leamer (2007) notes the housing market predicted eight of the ten post World War II recessions, acting as a leading indicator for the real sector of the economy. In fact, he goes as far as to state that "Housing is the Business cycle". More recently Balcilar *et al.*, (2014) provide evidence of the role of housing prices in causing even the "Great Depression." Recently, several authors indicate that house prices can be instrumental in forecasting output. (Forni *et al.*, 2003; Stock and Watson, 2003; Gupta and Das, 2010; Das *et al.*, 2009; 2010; 2011; Gupta and Hartley, 2013). The housing construction sector represents a large part of total economic activity expressed in the GDP. Consequently, as it reflects a large portion of the overall wealth of the economy, house prices fluctuations can be an indicator of the evolution of GDP (Case *et al.*, 2005). As it is the case with other assets, the movement of house prices can be also an indicator of the future direction of inflation as well (Gupta and Kabundi, 2010). Overall, accurate forecasting of the evolution path of house prices can be a useful tool both to housing market participants and monetary policy authorities.

There is a vast literature regarding U.S. house prices forecasting, both at regional and national levels. Rapach and Strauss (2007) use an autoregressive distributed lag (ARDL) model

* Corresponding author. Department of Economics, University of Pretoria, Pretoria, 0002, South Africa. Email: rangan.gupta@up.ac.za.

** Center of Advanced Statistics and Econometrics, Soochow University, China.

framework, containing 25 determinants to forecast real housing price growth for the individual states of the Federal Reserve's Eighth District. They find that ARDL models tend to outperform a benchmark AR model. Rapach and Strauss (2009) extend the same analysis on the 20 largest U.S. states based on ARDL models examining state, regional and national level variables. Once again, the authors reach similar conclusions on the importance of combining forecasts of models with different lag structure. Gogas and Pragidis (2010) use the risk premium calculated as the difference between various long-term interest rates and the agents' expectations about future short-term rates as an input variable in predicting the future direction of house prices. They conclude that investors and analysts can use effectively the information provided by the interest rate risk premium today in order to estimate the probability of obtaining a below-trend S&P CS-10 index three months ahead.

Gupta and Das (2010) also forecast the recent downturn in real house price growth rates for the twenty largest U.S. states. The authors use Spatial Bayesian VARs (BVARs), based only on monthly real house price growth rates, to forecast their downturn over the period 2007:01 to 2008:01. They find that BVAR models are well-equipped in forecasting the future direction of real house prices, though they significantly underestimate the decline. They attribute this under-prediction of the BVAR models to the lack of any information on fundamentals in the estimation process. Das *et al.*, (2010) use small-scale BVARs, Bayesian Factor Augmented VARs (BFAVARs) and large-scale BVARs in forecasting house prices of the nine census regions. The authors use the standard Minnesota Bayesian prior in estimating the Bayesian models. They indicate that the BFAVARs are best-suited in forecasting the house price growth rates of the nine census divisions. Gupta *et al.*, (2011) examine the explanatory power of small and large sets of economic variables, using atheoretical models such as VAR, BVAR, FAVAR, BVAR, BFAVAR, and forward-looking structural Dynamic Stochastic General Equilibrium (DSGE) models. Based on the average root mean-squared errors for the one-, two-, three-, and four-quarters-ahead forecasts, they find that the small-scale Bayesian-shrinkage model fed with 10 variables outperforms all the other models, however, the state-space-based DSGE model was the only model to predict the downturn. Similar observations, in terms of forecasting the downturn in house prices, have also been recently made by Zietz and Traian using state-space models as well. Gupta (2013) uses dynamic factor and Bayesian shrinkage models in a large number of predictors (145 variables) and forecasts house prices for four U.S. census regions and for the aggregate economy. The results show that the BFAVAR models exhibit the best forecasting ability. Superior performance of factor models in forecasting housing prices have also been depicted by Bork and Møller (2012). Similar results were also obtained by Gupta and Kabundi (2010) and Aye and Gupta (forthcoming) using Bayesian predictive regressions in forecasting the overall US house price index.

More recently, Balcilar *et al.*, (forthcoming) compared the ability of nonlinear AR models in forecasting nominal house price growth rates of the four U.S. census regions and the aggregate economy, relative to an AR model. Interestingly, even though they could detect nonlinearity in the in-sample for all the 5 growth rates of house prices, when it came to out-of-sample point, interval and density forecasting, the evidence in favor of the nonlinear model was virtually non-existent. These results were in sharp contrast to earlier evidence on forecasting house prices using univariate nonlinear, as in Miles (2008) and Cabrero *et al.*, (2011), but are similar to Barari *et al.*, (2014). While, Plakandaras *et al.*, (2014) combines Ensemble Empirical Mode Decomposition from the field of signal processing with the machine learning Support Vector Regression (SVR)

methodology for forecasting house prices. Their results point to superior performance of the SVR methodology when compared to Bayesian models. Similarly, Bork and Møller (forthcoming) indicates superior forecasting performances from time-varying, hence nonlinear, multivariate (based on regional and national predictors) models for state-level house price forecasting.

Clearly then, large number of models; linear and nonlinear, univariate and multivariate have been used in forecasting house prices, with results contingent upon not only the type of models chosen, predictors used and also periods under study.¹ Against this backdrop, the objective of this paper is to compare the ability of various univariate copula models, relative to standard benchmarks (random walk and autoregressive models) in forecasting real US house price over the annual out-of-sample period of 1874-2013, based on an in-sample of 1831-1873. Note that, since we work with real housing returns, we lose the first observation (1830) related to real house price. The starting point of the out-of-sample period, over which our model is recursively estimated, is driven by the fact that the first structural break, based on the Bai and Perron (2003) test, in the real housing returns is obtained in 1874. Our study is unique, not only because it is the first study to use copula models in forecasting house prices, but also because our study uses the longest available house price series for the US economy and hence, allows us to capture best the evolution of the US housing market over time. Further note that the decision to use an univariate approach is not only governed by the lack of standard predictors of house prices over this prolonged period, but primarily because, with house prices shown to be a leading indicator for the US economy, it is more pragmatic to obtain house price forecasts independent of its so-called predictors.

At this point, it is important to emphasize the reasons behind our decision to use copula models in forecasting house prices. First, copulas can be used to define nonparametric measures of dependence for pairs of random variables, which in our case happens to be the real housing price returns and its first lag. When fairly general and/or asymmetric modes of dependence are relevant, such as those that go beyond correlation or linear association, then copulas play a special role in developing additional concepts and measures. Also, statisticians or econometricians often possess more information about marginal distributions of related variables than their joint distribution. The copula approach is a useful method for deriving joint distributions given the marginal distributions, especially when the variables are non-normal. Finally, copulas are useful extensions and generalizations of approaches for modeling joint distributions and dependence that have appeared in the literature. As we show below based on statistical tests, not only does house price evolve in a nonlinear fashion,² but is also nonlinearly

¹ For a detailed literature review on forecasting involving the U.S. commercial and residential real-estate markets, refer to Ghysels *et al.*, (2013).

² Conventional wisdom argues that housing prices in the US rise more quickly and fully to market events that increase the equilibrium price than they do to market events that lower the equilibrium price. For example, the fall of housing prices during the recent financial crisis and Great Recession and beyond did not occur quickly enough to clear housing markets around the country, significantly slowing the recovery process. Recent studies (Genesove and Mayer, 2001; Engelhardt, 2001; Seslen, 2004; Kim and Bhattacharya, 2009; Balcilar *et al.*, 2011; Balcilar *et al.*, forthcoming) document evidence of such nonlinearity in housing prices. Several possible

related to its past, but real housing price returns are also non-normal; clearly then, making the variable under consideration suitable for being modeled using copulas. The rest of the paper is organized as follows: Section 2 discusses the basics behind copulas, with Section 3 presenting the data and results. Finally, Section 4 concludes the paper.

2. The Basics Copula Theory:

In probability theory and statistics, a copula is a joint probability distribution for which the marginal probability distribution of each variable is uniform. Copulas are popular in high-dimensional statistical applications as they allow one to easily model and estimate the distribution of random vectors by estimating marginals and copula separately. Copulas are used to describe the dependence between random variables, in our case the current and the lagged value of real housing returns. There are many parametric copula families available (as described below), which usually have parameters that control the strength of dependence.

The theorem underlying copulas was introduced in a 1959 article by Sklar written in French; a similar article written in English followed in 1973 (Sklar, 1973). The idea of copula had previously appeared in a number of papers, most notably in Hoeffding (1940, 1941) who established best possible bounds for these functions and studied measures of dependence that are invariant under strictly increasing transformations. Relationships of copulas to other work is described in Nelsen (2006).

Succinctly stated, copulas are functions that connect joint distributions to their one-dimensional margins. If F is an m -dimensional cumulative distribution function (cdf) with one-dimensional margins F_1, \dots, F_m , then there exists an m -dimensional copula C such that $F(y_1, \dots, y_m) = C(F_1(y_1), \dots, F_m(y_m))$.

Sklar's Theorem states that any joint joint distribution can be written in terms of univariate marginal distribution functions and a copula which describes the dependence structure between the variables.

explanations for intrinsic nonlinearity in house prices exist. First, as noted above, households respond asymmetrically over the business cycle. Abelson *et al.* (2005) argue that households more likely buy when prices rise, because they expect further rises and try to avoid higher payments. Households will less likely buy or sell, however, due to loss aversion with falling house prices. Seslen (2004) argues that households exhibit forward-looking behavior and a higher probability of trading up, during expansions, since equity constraints prove less binding. During the downswing of the housing market cycle, households less likely trade, implying downward rigidity of house prices. Loss aversion during the downswing more likely reduces the mobility of households as well as trading activity. Further, Muellbauer and Murphy (1997) note that the presence of lumpy transaction costs in the housing market can also cause non-linearity. Given these issues, it makes sense to test for non-linear housing price movements.

Some examples of popular copulas are Gaussian and t copulas. A particular group of copulas that has proved useful in empirical modeling is the Archimedean class. Archimedean copulas are popular because they are easily derived and are capable of capturing wide ranges of dependence. Some examples in the Archimedean class are the Gumbel, Frank, Joe, Clayton, Ali Mikhail Huq. In this paper, we considered the following types of copulas:

Gaussian (Normal) copula: The normal copula takes the form

$C(u_1, u_2; \theta) = \Phi_G(\Phi^{-1}(u_1), \Phi^{-1}(u_2); \rho)$, where Φ is the cumulative distribution function of the standard normal distribution, and $\Phi_G(u_1, u_2; \rho)$ is the standard bivariate normal distribution with correlation parameter ρ restricted to the interval $(-1, 1)$. The value of $\rho=0$, corresponds to independence.

Student's t -copula: An example of a copula with two dependence parameters is that for the bivariate t -distribution with ν degrees of freedom and correlation ρ , takes the form

$C^t(u_1, u_2; \rho) = T_G(T^{-1}(u_1), T^{-1}(u_2); \rho)$, where T is the cumulative distribution function of the t distribution with ν degree of freedom, and $T_G(u_1, u_2; \rho)$ is the bivariate t distribution with ν degree of freedom with correlation parameter ρ restricted to the interval $(-1, 1)$. The value of $\rho=0$, like the Gaussian copula, corresponds to independence.

The Clayton copula (Clayton, 1978) copula, also referred to as the Cook and Johnson (1981) copula, originally studied by Kimeldorf and Sampson (1975), takes the form:

$C(u_1, u_2; \rho) = (u_1^{-\rho} + u_2^{-\rho} - 1)^{-1/\rho}$ with the dependence parameter ρ restricted on the region $(0, \infty)$. As ρ approaches zero, the marginals become independent. The Clayton copula cannot account for negative dependence.

The Frank copula (1979) takes the form:

$C(u_1, u_2; \rho) = -\rho^{-1} \{ \log 1 + (e^{-\rho u_1} - 1)(e^{-\rho u_2} - 1) / (e^{-\rho} - 1) \}$, the dependence parameter may assume any real value $(-\infty, \infty)$. Value 0 corresponds to independence.

The Gumbel copula (1960) takes the form:

$C(u_1, u_2; \rho) = \exp -(\{-\log(u_1)\}^\rho + \{-\log(u_2)\}^\rho)^{1/\rho}$,

The dependence parameter ρ is restricted to the interval $[1, \infty)$. Value of 1 corresponds to independence.

The Joe copula (Nelson, 2006) takes the form:

$C(u_1, u_2; \rho) = 1 - \{ (1-u_1)^\rho + (1-u_2)^\rho - (1-u_1)^\rho (1-u_2)^\rho \}^\rho$. The dependence parameter ρ is restricted to the interval $[1, \infty)$. Value of 1 corresponds to independence.

The **Ali-Mikhail-Huq (AMH) copula** (Nelson, 2006) takes the form

$C(u_1, u_2; \rho) = u_1 u_2 / \{1 - \rho(1-u_1)(1-u_2)\}$, the dependence parameter ρ is restricted to the interval $[-1, 1)$. Value of 0 corresponds to independence.

3. Data and Results

The data examined correspond to inflation-adjusted, (i.e., real) house price returns for the aggregate US economy covering the period of 1831-2013. The real returns are computed as percentage change in the natural logarithmic values of the real house price, which in turn starts from 1830. Since, we work with returns, we lose the first observation. Our decision to work with returns, rather than levels, emanates from the need to use stationary data. The unit root tests have been reported in Table A1 of the Appendix. The data is sourced from Winans International, and the index measures real values of price of new homes going back to 1830. Figures 1a and 1b displays the (natural) logged time series and its corresponding returns (in percentage) version, respectively. Table 1 provides a set of summary statistics for the real house price returns. As can be seen, the Jarque-Bera test rejects the null of normality and hence, provides one of the motivation to use Copula. Further, when Brock et al., (1996; BDS) test of nonlinearity is performed, as reported in Table 2, on the real rates of return, the residuals obtained from a regression involving the real rates of return with just a constant (Naive), and the residuals from an autoregressive model of order one (AR(1)) for the real rates of return, we overwhelmingly reject the null hypothesis that the series are *i.i.d.*, implying nonlinearity in the data generating process. Consequently, these results involving non-normality and nonlinearity, motivate our use of Copula models.

Figure 1a. Real House Price

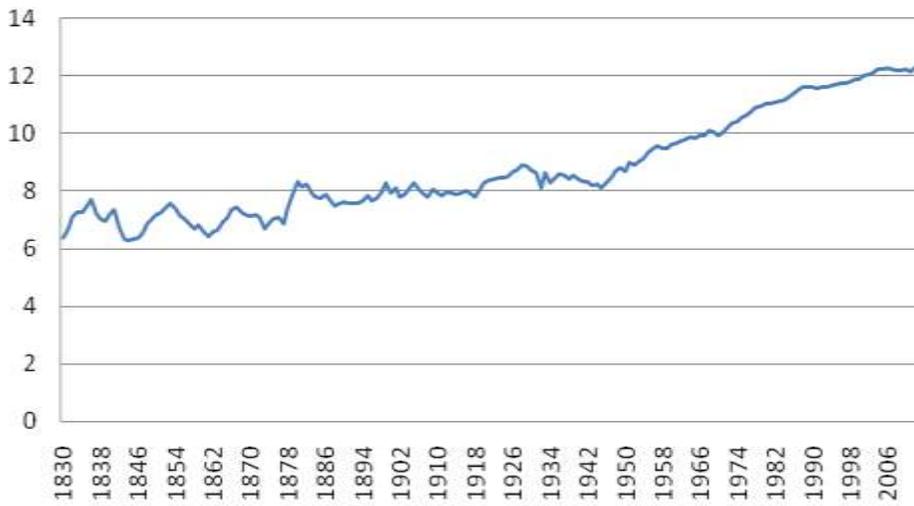


Figure 1b. Real House Price Return

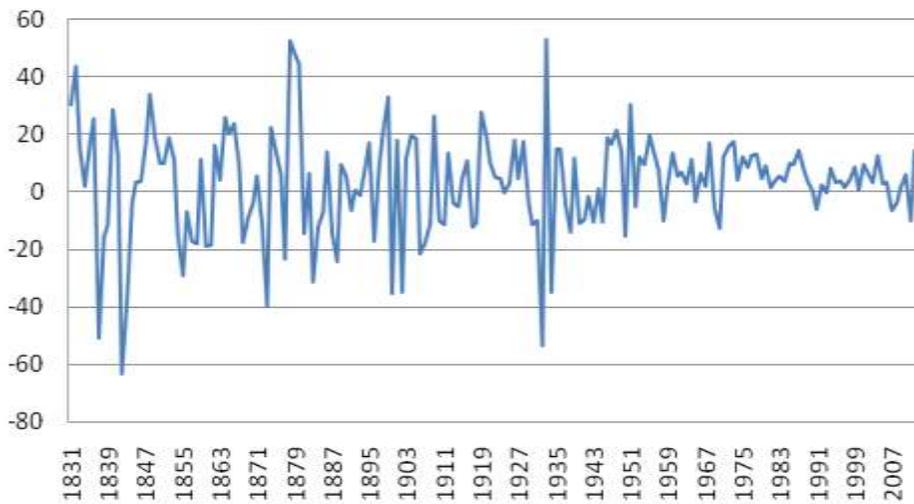


Table 1. Summary Statistics for the Real House Price Returns: (1831-2013)

Observations	183
Mean	3.28
Median	4.68
Maximum	53.03
Minimum	-63.23
Std. Dev.	17.80
Skewness	-0.49
Kurtosis	4.72
Jarque-Bera	29.85***

Notes: Std. Dev.: Standard Deviation; *** indicates the rejection of the null of normality at 1 percent level of significance.

Table 2. BDS (1996) Test of Nonlinearity

<i>Dim</i>	Real House Price Returns	RW Residuals	AR (1) Residuals
2	0.0000	0.0000	0.0000
3	0.0000	0.0000	0.0000
4	0.0000	0.0000	0.0000
5	0.0000	0.0000	0.0000
6	0.0000	0.0000	0.0000

Notes: *Dim* stands for dimension. Table reports the p-values of the *BDS* test.

To decide on the length of our in-sample and out-of-sample, we carried out direct tests of multiple structural breaks on the real house price returns using the Bai and Perron (2003) tests. It must be pointed out that, we estimate the various copula models as well as the benchmark Naive and AR(1) models recursively over the out-of-sample period. Based on the Bai and Perron (2003) tests of structural breaks, with a maximum number of breaks set to 5 and an end point trimming of 5% of the observations, we detected 5 breaks for the real house price returns series (regressed on a constant and its first lag), at the following dates: 1874, 1883, 1920, 1933, 1946. The break point tests have been reported in Table A2 of the Appendix. Note that, even if there is additional structural breaks within the 5 % trimming areas, for instance the recent subprime crisis, they would be captured by the nonlinear modeling of the copula within the in-sample and the recursive estimation of all the models over the out-of-sample. Given that the first

break occurred in 1874, our full-sample is split into 1831-1873 as the in-sample and the 1874-2013 as the out-of-sample.

Table 3 reports the in-sample and out-of-sample predictive ability of the models, based on the Akaike Information Criteria (AIC) and the mean squared error (MSE) metric, respectively. Turning first to the in-sample fit of the model, the Student's t copula³ produces the lowest AIC value followed by Normal copula. The fits of the rest of the copula models are virtually the same. The AR(1) and the Naive model produces the worst fit, with the latter coming in last in the rankings. When we focus on out-of-sample forecasting of one year-ahead forecasts, the results are slightly different. Again, as with the in-sample results, the Student's t copula produces the lowest MSE, followed by the Normal Copula. Barring the AMH copula which produces the highest MSEs, it is difficult to distinguish between the Clayton, Frank, Gumbel and Joe Copula models. Again, the Naive and AR(1) models perform poorly in comparison to all the copula models in terms of the MSEs. However, unlike in terms of the AIC for the in-sample fit, the AR(1) model is now ranked below the Naïve model. Further, based on the, the Diebold and Mariano (1995) test statistic of equal forecast accuracy, the Student's t -copula model not only significantly outperforms the Naive and the AR(1) model, but also the next-best performing Normal copula model at one percent level of significance. All in all, our results provide overwhelming evidence in favor of the copula models, and especially the Student's t copula, both in terms of in-sample and out-of-sample predictability results.

Table 3. In-sample and Out-of-sample Forecasting Results

Models	AIC (1831-1873)	MSE (1874-2013)
Naive	9.36	257.78
AR(1)	9.24	273.81
Gaussian (Normal) copula	1.45	31.63
Student's t copula	0.60	24.97 {-5.67***}[-6.21***](-3.33***)
Clayton copula	2.00	68.57
Frank copula	2.01	68.56
Gumbel copula	2.00	68.56
Joe copula	2.00	68.56
AMH copula	2.00	217.48

³ The best-fit student's t copula model had degrees of freedom equal to 20.

Notes: AIC: Akaike Information; MSE: Mean squared error; AMH: Ali-Mikhail-Huq copula; Bold-italic entries corresponds to the minimum values for the AIC and the MSE corresponding to a specific model; **{***}***** indicates that the Student's t copula significantly outperforms the {RW} [AR(1)] (Gaussian (Normal) copula) model at 1% level of significance based on Diebold and Mariano (1995) test of equal forecast accuracy. The corresponding p -values were: {1.424E-08}[5.39E-10](0.0009).

4. Conclusions

The recent global financial crisis has resulted in a keen interest in both academic and policy circles on the role of asset prices, and in particular housing prices, on economic activity. Also, there exists evidence that house prices lead inflation and output of the US economy. Hence, accurate forecasting of the evolution path of house prices can be a useful tool both to housing market participants and monetary policy authorities. There is a vast literature regarding U.S. house prices forecasting, both at regional and national levels, with large number of models; linear and nonlinear, univariate and multivariate frameworks used in forecasting house prices, with results contingent upon not only the type of models chosen, predictors used and also periods under study.

Against this backdrop, the objective of this paper is to compare the ability of various univariate copula models, relative to standard benchmarks (naive and autoregressive models) in forecasting real US house price over the annual out-of-sample period of 1859-2013, based on an in-sample of 1831-1873. Our study is unique, not only because it is the first study to use copula models in forecasting house prices, but also because our study uses the longest available house price series for the US economy. Our decision to use a wide variety of copula models, namely Normal, Student's t , Clayton, Frank, Gumbel, Joe and Ali-Mikhail-Huq, is to account for non-normality and nonlinearity in the data generating process of the real housing price returns. Overall, our results provide overwhelming evidence in favor of the copula models relative to linear benchmarks, and especially for the Student's t copula, which outperforms all other models both in terms of in-sample and out-of-sample predictability results. Our results highlight the importance of accounting for non-normality and nonlinearity in the data generating process of real house price returns for the US economy for nearly two centuries of data.

References

- Abelson, P., Joyeux, R., Milunovich, G. and Chung, D. (2005). Explaining house prices in Australia: 1970–2003. *The Economic Record*, 81 (S1), S96-S103.
- Aye, G. and Gupta R. (forthcoming). Forecasting real house price of the US: an analysis covering 1890 to 2012. *Journal of Economic Computation and Economic Cybernetics Studies and Research*.
- Bai J., and Perron P. (2003). Computation and Analysis of Multiple Structural Change Models. *Journal of Applied Econometrics*, 18, 1–22.
- Balcilar, M., Gupta, R., and Shah, Z. B. (2011). An in-sample and out-of-sample empirical investigation of the nonlinearity in house prices of South Africa. *Economic Modelling* 28, 891-899.

- Balcilar, M., Gupta, R. and Miller, S.M. (2014). Housing and the great depression. *Applied Economics*, 46, 2966-2981.
- Balcilar, M., Gupta, R. and Miller, S.M. (forthcoming). The out-of-sample forecasting performance of non-linear models of regional housing prices in the US. *Applied Financial Economics*.
- Barari, M., Sarkar, N., Kundu, S., and Chowdhury, K. B. (2014). Forecasting house prices in the United States with multiple structural breaks. *International Econometric Review* 6, 1-23.
- Bork, L. and Møller, S. V. (forthcoming). Forecasting house prices in the 50 states using Dynamic Model Averaging and Dynamic Model Selection. *International Journal of Forecasting*.
- Bork, L. and Møller, S. V. (2012). Housing price forecastability: A factor analysis. *EFA 2012 Copenhagen Meetings Paper*.
- Brock, W., D. Dechert, J. Scheinkman and LeBaron, B. (1996). A test for independence based on the correlation dimension. *Econometric Reviews*, 15, 197–235.
- Cabrera, J. F., Wang, T., and Yang, J. (2011). Linear and nonlinear predictability of international securitized real estate returns: A reality check. *Journal of Real Estate Research* 33, 565-594.
- Case, K., Shiller, R., and Quigley, J. (2005). Comparing wealth effects: The stock market versus the housing market. *Advances in Macroeconomics* 5, 1-32.
- Clayton, D. G. (1978). A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika*, 65, 141–151.
- Cook, R. D. and Johnson, M.E. (1981). A family of distributions for modelling non-elliptically symmetric multivariate data. *Journal of Royal Statistical Society B* 43(2), 210–218. Berlin and New York: Springer-Verlag.
- Das, S., Gupta, R. and Kabundi, A. (2009). Could we have predicted the recent downturn in the South African housing market? *Journal of Housing Economics*, 18 (4), 325–335.
- Das, S., Gupta, R. and Kabundi, A. (2010). The blessing of dimensionality in forecasting real house price growth in the nine census divisions of the US. *Journal of Housing Research*, 19 (1), 89–109.
- Das, S., Gupta, R. and Kabundi, A. (2011). Forecasting regional house price inflation: a comparison between dynamic factor models and vector autoregressive models. *Journal of Forecasting*, 30 (2), 288–302.
- Diebold, F. X., and Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business and Economics Statistics* 13, 253-263.
- Engelhardt, G. V. (2001). Nominal loss aversion, housing equity constraints, and household mobility: Evidence from the United States. *Center for Policy Research Working Paper*, No. 42, 1–61.
- Forni M, Hallin, M, Lippi, M, and Reichlin, L. (2003). Do financial variables help forecasting inflation and real activity in the euro area? *Journal of Monetary Economics* 50, 1243-1255.
- Frank, M. J. (1979). On the simultaneous associativity of $F(x,y)$ and $x+y - F(x,y)$. *Aequationes Math* 19, 194–226.
- Genesove, D., and Mayer, C. J. (2001). Nominal loss aversion and seller behavior: Evidence from the housing market. *Quarterly Journal of Economics* 116, 1233–1260.

- Ghysels E., Plazzi A., Torous W. and Valkanov R. (2013). Forecasting Real Estate Prices, E., Alberto, W., and R. Valkanov, *Handbook of Economic Forecasting: Vol II*, G. Elliott and A. Timmermann (Eds.), Elsevier.
- Gogas, P. and Pragidis, I. (2010). Does the interest risk premium predict housing prices?, *DUTH Research Papers in Economics 1-2010*, Democritus University of Thrace, Department of Economics.
- Gumbel, E. J. (1960). Distributions des Valeurs Extremes en Plusieurs Dimensions. *Publications de l'Institut de Statistique de l'Université de Paris* 9, 171–173.
- Gupta, R and Kabundi, A. and Miller, S.M. (2011). Forecasting the US real house price index: structural and non-structural models with and without fundamentals. *Economic Modelling*, 28 2013–2021.
- Gupta, R. (2013). Forecasting house prices for the four census regions and the aggregate US economy in a data-rich environment, *Applied Economics*, 45:33, 4677-4697.
- Gupta, R., and Das, S. (2010). Predicting downturns in the US housing market: A Bayesian approach. *The Journal of Real Estate Finance and Economics* 41, 294-319.
- Gupta, R. and Hartley, F. (2013). The role of asset prices in forecasting inflation and output in South Africa. *Journal of Emerging Market Finance*, vol. 12(3), pp. 239- 291.
- Gupta, R. and Kabundi, A. (2010). Forecasting real US house price: principal components versus Bayesian regressions. *International Business and Economics Research Journal*, 9 (7), 141–152.
- Hoeffding, W. (1940). Scale-invariant correlation theory. In: N. I. Fisher and P. K. Sen (eds.): *The Collected Works of Wassily Hoeffding*. New York: Springer-Verlag, pp. 57–107.
- Hoeffding, W. (1941). Scale-invariant correlation measures for discontinuous distributions. In: N. I. Fisher and P. K. Sen (eds.): *The Collected Works of Wassily Hoeffding*. New York: Springer-Verlag, pp. 109–133.
- Kim, S-W., and Bhattacharya. R. (2009). Regional housing prices in the USA: An empirical investigation of nonlinearity. *Journal of Real Estate Finance and Economics* 38, 443-460.
- Kimeldorf, G. and Sampson, A. R. (1975). Uniform representations of bivariate distributions. *Communications in Statistics* 4, 617–627.
- Leamer, E. E. (2007). Housing is the business cycle. In *Housing, Housing Finance, and Monetary Policy*, Economic Symposium Conference Proceedings, Kansas City Federal Reserve Bank, 149-233.
- Miles, W. (2008). Boom-Bust cycles and the forecasting performance of linear and non-linear models of house prices. *Journal of Real Estate Finance and Economics* 36, 249-264.
- Muellbauer, J., & Murphy, A. (1997). Booms and busts in the UK housing market. *Economic Journal* 107, 701–727.
- Nelsen, R. B. (2006). *An Introduction to Copulas*. 2nd edition. New York: Springer.
- Plakandaras, V., Gupta, R., Gogas, P., and Papadimitriou, T. (2014). Forecasting the U.S. real house price index. *University of Pretoria, Department of Economics, Working paper No. 2014-18*.
- Rapach, D.E. and Strauss, J.K. (2007), Forecasting real housing price growth in the eighth district states. Federal Reserve Bank of St. Louis, *Regional Economic Development*, 3, 33–42.
- Rapach, D. E. Strauss, J. K. (2009), Differences in housing price forecastability across US states. *International Journal of Forecasting*, 25, 351–372.

Seslen, T. (2004). Housing price dynamics and household mobility decisions. Mimeo, University of Southern California.

Sklar, A. (1973), Random variables, joint distributions, and copulas. *Kybernetika* 9, 449–460.

Stock, J. H., and Watson, M. W. (2003). Forecasting output and inflation: The role of asset prices. *Journal of Economic Literature* 41, 788-829.

Zietz, J., and Traian, A. (2014). When was the U.S. housing downturn predictable? A comparison of univariate forecasting methods. *The Quarterly Review of Economics and Finance* 54, 271-281.

APPENDIX

Table A2. Multiple Break Points

Tests

Bai-Perron (2003) tests of 1 to M globally determined breaks

Sample: 1830 2013

Included observations: 182

Breakpoint variables: C RRHP(-1)

Break test options: Trimming 0.05, Max. breaks 5, Sig. level 0.10

Allow heterogeneous error distributions across breaks

Sequential F-statistic determined breaks:				2
Significant F-statistic largest breaks:				5
UDmax determined breaks:				1
WDmax determined breaks:				4
Breaks	F-statistic	Scaled F-statistic	Weighted F-statistic	Critical Value
1 *	6.314863	12.62973	12.62973	11.02
2 *	5.349608	10.69922	11.25051	10.48
3	4.432756	8.865511	10.16628	9.61
4 *	5.975551	11.95110	14.64974	8.99
5 *	5.008740	10.01748	12.98737	8.50
UDMax statistic*		12.62973	Critical value**	11.69
WDMax statistic*		14.64974	Critical value**	12.33

* Significant at the 0.10 level.

** Bai-Perron (Econometric Journal, 2003) critical values.

Estimated break dates:

1: 1881

2: 1881, 1946

3: 1881, 1930, 1946

4: 1881, 1920, 1933, 1946

5: 1874, 1883, 1920, 1933, 1946