İzmir İktisat Dergisi
İzmir Journal of Economics

ISSN:1308-8173 **E-ISSN:** 1308-8505 **YIL:** 2021 **Cilt:** 36 **Sayı:** 2 **Sayfa:** 259-273
Geliş Tarihi: 23.05.2020 **Kabul Tarihi:** 29.07.2021 **Online Yayın:** 05.08.2021 **Doi:** 10.24988/ije.202136202

ÖZGÜN ARAŞTIRMA

The Impact of Macro-Economic Drivers in Housing Markets: The US Case

Bilgi YILMAZ [1] , Fatma YERLİKAYA ÖZKURT [2] , A. Sevtap SELCUK-KESTEL [3]

Abstract

This paper analyzes the effect of macro-economic, financial and commodity market indicators on housing markets. We compare the efficiency of the models generated by Generalized Linear Models (GLM) and Multivariate Adaptive Regression Splines (MARS) according to method free measures for estimating the housing market trend. These models are used for the first time to identify the influence of macro-economic indicators on housing markets and the estimation of the trend in housing markets to our best knowledge. The empirical analysis focuses on the US housing market, and the illustration of the proposed models is done through the monthly historical realizations of S\&P/Case-Shiller National Home Price Index (HPI) and the US macro-economic indicators over the period from 1999-January to 2018-June. It contributes to the literature by highlighting the interaction between macro-economic indicators and housing markets and analyzing the mechanism of housing markets. The findings indicate that the house price trends are estimated with more accuracy and these models capture the joint influence of explanatory variables. Further, the MARS method is shown to outperform GLM compared to the prediction and forecasting power.

Keywords: Housing Market, Generalized Linear Models, Multivariate Adaptive Regression Splines
Jel Kodu: C01, C53, C44, C63

Konut Piyasalarında Makroekonomik Faktörlerin Etkisi: ABD Örneği

Özet

Bu makale makroekonomik, finansal ve emtia piyasaları göstergelerinin konut piyasaları üzerindeki etkisini analiz etmektedir. Genelleştirilmiş Doğrusal Modeller (GDM) ve Çok Değişkenli Uyarlanabilir Regresyon Eğrileri (ÇDRE) tarafından üretilen modellerin yeterliliğini konut piyasası eğilimini tahmin etmek için bağımsız ölçüm yöntemlerine göre karşılaştırıyoruz. Araştırmalarımıza göre bu modeller ilk kez makroekonomik göstergelerin konut piyasaları üzerindeki etkisini ve konut piyasalarındaki eğilimine yönelik tahmini belirlemekte kullanılmaktadır. Ampirik çalışmalar, ABD konut piyasalarına odaklanmakta ve önerilen modellerin gösterimi Ocak 1999-Haziran 2018 periyodu arasında gözlemlenen aylık S\&P/Case-Shiller Ulusal Konut Fiyat İndeksine ve ABD macroeconomic faktörlerine uygulanmaktadır. Bu çalışma makro ekonomik göstergeler ve konut piyasaları arasındaki etkileşimi vurgulayarak ve konut piyasalarının mekanizmasını analiz ederek literatüre katkıda bulunmaktadır. Bulguları, konut fiyat eğilimlerinin daha doğru bir şekilde tahmin edildiğini ve bu modellerin açıklayıcı değişkenlerin ortak etkisini yakaladığını göstermektedir. Ayrıca, ÇDRE yönteminin tahmin ve geleceğe yönelik tahmin gücüne kıyasla GDM'den daha iyi performans gösterdiğini ortaya koymuştur.

Anahtar kelimeler: Konut Piyasası, Genelleştirilmiş Doğrusal Modeller, Çok Değişkenli Uyarlanabilir Regresyon Eğrileri
Jel Kodu: C01, C53, C44, C63

1. INTRODUCTION

The real estate market is one of the leading and locomotive markets of national economies due to its high dependence on domestic capital, its creation of high added value, the magnitude of its employment potential, and its strong relation to other markets, such as the financial and commodity markets. In the last two decades, increase in the acceleration of capital flows across countries, thanks to the economic globalization and the addition of property-based investment tools to the field at which the capital flows are interested in has prompted the

ATIF ÖNERİSİ (APA): Yilmaz, B.; Yerlikaya Özkurt, F., Selcuk-Kestel, A. S. (2021). The Impact of Macro-Economic Drivers in Housing Markets: The US Case. İzmir İktisat Dergisi. 36(2). 259-273. Doi: 10.24988/ije.202136202

[1] Post.Doc. Researcher. Middle East Technical University, Institute of Applied Mathematics, Çankaya/ANKARA.
EMAIL: bilgiyilmaz07@gmail.com **ORCID:** 0000-0002-9646-2757
[2] Assist. Prof. Dr. Atılım University, Department of Industrial Engineering, Gölbaşı /ANKARA,
EMAIL: fatma.yerlikaya@atilim.edu.tr **ORCID:**0000-0002-4747-8461
[3] Prof. Dr. Middle East Technical University, Institute of Applied Mathematics, Çankaya/ANKARA,
EMAIL: skestel@metu.edu.tr **ORCID:** 0000-0001-5647-7973

impact of real estate markets, especially housing markets, on national economies.

As assets in housing markets constitute a considerable amount of households' wealth and Gross Domestic Product (GDP), they naturally shape long-term developments of economies. In this respect, housing markets play a crucial role in national economic activities. For instance, crashes of housing bubbles generally destabilize economies, and more importantly, the crashes cause significant economic recessions, which may lead to global financial crises like the one in 2008. Therefore, the state of the current economy and the latest global financial crisis led researchers to pay more attention to the role of housing markets on national economies (Valentini et al., 2013).

In the last two decades, there exists vast amount of studies in the literature that emphasize the relationship between housing markets and macro-economic indicators (Agnello and Schuknecht, 2011; Aspachs-Bracons and Rabanal, 2010; Iacoviello, 2005; Lambertini et al. 2017). These studies show that interest rate, mortgage rate, inflation, unemployment rate, industrial production, and financial indicators, like the exchange rate inevitably influence households in consumption and investment decisions. In return, they affect housing markets with a lag depending on propagation mechanism's speed. The propagation speed strongly influenced by the efficiency of the institutional framework, such as land availability, zoning regulations, and the rate of administrative processes (Adams and Füss, 2010). Alongside these indicators, it is known that credit supply, transaction costs, and innovations in mortgage products also have a significant role in housing markets. For instance, if changes in spot rates affect mortgage rates quickly, an increase in the money supply influences housing markets faster than the fixed mortgage rate case.

Demographic changes also have a significant effect on housing markets and households' behaviour. For instance, Mankiw and Weil (1989) suggest that the decline in the fraction of the US population in the prime house-buying age bracket during the 1990s caused a substantial fall in actual house prices. However, household behaviours are more sensitive to the movements in interest rates, particularly if they are unexpected, and changes in household income particularly caused by unemployment. The sensitivity of household behaviours to interest rate changes depends on whether the interest rate on the debt is predominantly fixed or variable over the life of the mortgage loan. In this regard, Poterba (1991) finds that changes in borrower costs associated with interest rate movements and tax policy are an essential determinant of real house price movements. It determines whether households, financial intermediaries, or pension funds are mostly exposed to changes in interest rates based on the location. In turn, this will influence the short-term impact of changes in the interest rate.

In Johness and Hyclak (1999), the relationship between the labour income and house price is investigated, and it finds some evidence that changes in unemployment affect house prices. Moreover, in empirical studies, such as Englund and Ionnides (1997), Malpezzi (1999), Poterba (1991), it is proved that income is one of the most significant drivers of house prices. Most of these studies rely on average income measures, such as per capita disposable income. Such average price measures of housing markets capture the fact that the wealthier households demand more consumption good and thus more housing than poorer households (Ortalo-Magne and Raady, 2006).

There are also studies in the literature, which identify house prices, displaying a feedback reaction to economies. For instance, an increase in house prices makes house-owners feel wealthier because of their collateral size and the value of the house (Adams and Füss, 2010). Furthermore, if a house-owner has a liquidity constraint, the increase in his/her house price may be his/her only opportunity to borrow. This kind of wealth shocks causes an increase in households' consumption. On the other hand, if house prices decline, it leads to an adverse impact on households' consumption. It is because a decrease in house prices increases the number of mortgage defaults, which reduces the supply of bank credit as banks lose part of their capital (Glaeser and Parker, 2000).

Even though there exist studies that investigate housing markets, analysing them is more complicated and cumbersome than fully competitive markets since house prices do not respond to economic fluctuations as fast as in fully competitive markets. Generally, house prices show a steady downward price movement since house-owners resist selling their houses under a specific price barrier during recessions. As a result, house prices have a decreasing trend through high inflationary periods rather than through formal price reduction (Adams and Füss, 2010). Besides, the price inertia also affects the house price behaviour during economic booms since high expectations of house-owners facilitate housing bubbles. Furthermore, along with such problems, housing markets have specific characteristics distinguished from fully competitive markets. First of all, housing markets are highly illiquid due to high transaction costs and the time to spend the decision of a house to purchase. Generally, real prices are known only by the buyer and seller, which prevents the market from observing house prices. Entering and retreating housing markets are relatively tricky since they require a significant amount of cost. Moreover, housing markets are highly heterogeneous since houses are unmovable and stick to a location.

Despite the difficulties, it is vital to specify underlying drivers of house prices and their implications for housing markets. Therefore, the strong dependence on housing markets and economies allows us to construct statistical models, which capture the behaviour of housing markets as well as its underlying factors by using the celebrated Generalized Linear Models (GLM) and Multivariate Adaptive Regression Splines (MARS) methods and determine the effect of macro-economic factors on housing markets. Hence, unlike many other studies, this paper investigates the effect of macro-economic indicators on house prices by using linear and nonlinear parametric and nonparametric regression models. The study examines the influence of relevant indicators by using the monthly US market data over the last 19 years. In the proposed models as its rigorous mathematical descriptions include the influence of the historical prices as well as the impact of macro-economic indicators. S\&P/Case-Shiller National Home Price Index is analysed under the influence of Consumer Price Index, Civilian Unemployment Rate, 10-Year Treasury Constant Maturity Rate, 30-Year Fixed Rate Mortgage Average in the US, US/Euro Foreign Exchange Rate, Effective Federal Funds Rate, Working-Age Population (Aged 15-64 All Persons), Crude Oil Prices, Real Disposable Personal Income (Per Capita), and Recession Cycle in the Economy within the given period. The comparison of models resembling a better fit is made through their accuracy with method free error measures.

The organization of the paper is as follows: Section 2 presents a brief explanation of the US housing market. In Section 3, we briefly define different regression-based models that one being parametric (GLM) and the other being nonparametric (MARS), which is followed by macro-economic indicators taken into account and the empirical findings in Section 4. Section 5 concludes the paper.

2. THE US HOUSING MARKET

The US housing market has experienced a high degree of volatility cycle relative to economic indicators, such as consumer price index and real income levels, during the period 1998-2009 due to significant structural changes and fluctuations in the US economy (Guirguis et al., 2005). As it is well known, the 30-year fixed mortgage rate has dominated the US housing market within this period. Thus, the mortgage rate and spot rates are some of the essential features in explaining the US housing market behaviour. Figure 1 also reflects the dominance of the 30-year fixed mortgage. As external funding is mostly provided through mortgages in housing markets. There exists a high association between housing markets and both mortgage and spot rates. In Figure 1, by considering monthly 30-year fixed mortgage rate (FRM), 6-Month Treasury Bill, Effective Federal Funds Rate in the US, and S\&P/Case-Shiller National Home Price Index (HPI) series between 1975 and 2016, we observe that spot rates and mortgage rate follow similar pattern at which mortgage rate yield is higher than spot rates. It is an expected result as the treasury notes are the safest investment instruments since the US government issues guarantees on them. We also observe that both mortgage and spot rates show a reverse pattern and negative association to HPI. We observe two striking dynamic structures: i) The periodic patterns and opposite trend components are consistent features of the housing market. ii) There is a sharp rise, even during the 2001 recession, reaching a remarkable increase in 2006 compared to the position of mortgage rate, which is commonly perceived as a bubble and explained in detail in (Diewert et al., 2009). It is important to note that an increase in the house prices, even resembling a bubble, is triggered by the preceding house prices in time. Also, the mortgage rate and spot rates appear less vulnerable to financial crises compared to the house prices.

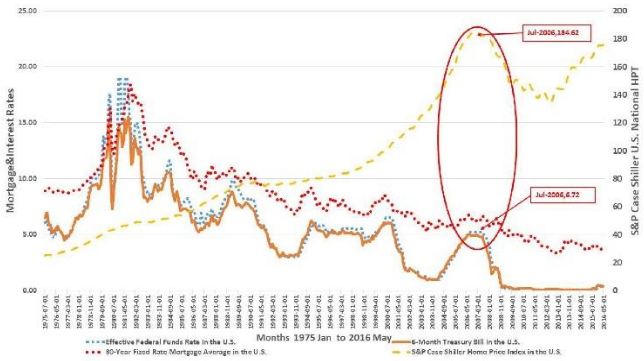

Figure 1: The development of the HPI, FRM, and interest rates between years 1975 and 2016 in the US

However, in the mid-2000s, non-traditional mortgage products, which allow consumers more straightforward access to credit, challenged the dominance of the 30-year fixed mortgage rate (Dokko et al., 2015). The consumers who are seeking a house may use the non-traditional mortgage products to purchase more expensive houses than they could afford with the expectation of a rise in house prices. Such use of alternative products in purchasing houses may cause negative externalities through the speculations (Gadi Barlevy, 2011). As we witnessed, during the period 2006-2009, housing and financial crisis elevated mortgage delinquencies, and defaults dampened the house prices, increased the pessimism among consumers and investors, and eventually ruined the US financial markets and spread to financial markets worldwide (Chauvet et al., 2016; Rapach and Straus, 2009). Within the period, the value of the US housing stock decrease $4.4 trillion, which is widely believed to be the primary reason behind the latest subsequent financial crises and recession (Carson and Dastrup, 2013). Almost all countries experienced the massive global financial crisis over the period 2006-2009 that originated from the demand in the US housing market.

The housing demand is also a significant indicator of housing markets. Figure 2 displays the number of privately owned new housing units starting each year (solid line), and recessions as determined by the National Bureau of Economic Research (shaded areas) in the US for the period January 1959 - July 2018. The demand trend helps to clarify the state of the economy. That is to say, it is typical for a significant decline in housing demand preceded by a recession starts to precede a recession. Comparing substantial recessions in the US economy, it is evident that a recession typically occurred soon after a peak in house prices. For instance, the number of privately-owned housing units peaked in 2006, and followed by a very steep decline, which overlaid with the recession period that we experienced in 2008. Balloon and bust business cycles related to house prices are often associated with similar behaviour in private credit. Thus, the reductions in housing demand might be the best early warning sign of an oncoming recession (Leamer, 2013). In this respect, the importance of the US housing market in the US economy is exposed in Figure 2.

Figure 2: The recession (shaded vertical lines) and the number of privately owned new houses (solid line)

Both Figure 1 and 2 also illustrate that the US housing market is about to finish its recovery period. There are four severe aspects behind the recovery of the US housing market that need to be considered. First, unlike the speculative increase in housing prices before the crisis, current house prices rise due to the fundamental strength of the US economy. Second, the US population growth and the increase in the housing demand. The rise in US citizens' wealth, namely, the increase in gross income per household after the financial crisis holds the third aspect. Fourth, the labour market's awakening during the recovery period in the US economy.

3. BASICS ON GLM AND MARS

3.1 The Generalized Linear Models (GLM)

Generalized Linear Models (GLM) provide researchers with a unified framework for multi-factor regression analysis by allowing the use of multiple regression, logistic regression, variance, and covariance analysis. These models consist of three components; a link function, g(.), that specifies the transformation of the response variable to be modelled by a linear function of the explanatory variables, an error distribution that is suitable for each type of response, and a variance function that specifies the relationship between the mean and variance of the error distribution. In GLM, normality and constant variance assumptions are no longer a requirement for the error component (Nelder and Beker, 1972).

Given a random variable vector, $\boldsymbol{Y} = (Y_1, Y_2, \ldots, Y_N)$, having a mean vector μ and a matrix $\boldsymbol{X}$ of order $N \times p$, and a p-dimensional parameter vector $\boldsymbol{\beta}$, a GLM has the primary objective to investigate the relationship between $\mu = \mathbb{E}[Y]$ and $\boldsymbol{X}$ through $\boldsymbol{\beta}$. Here, the parameter μ represents the systematic part of the model and it can be written by means the existence of covariates $x_1, x_2, \ldots, x_p$ and parameters β_j $(j = 1,2, \ldots, p)$ as

$$\mathbb{E}[Y_i] = \mu_i = \sum_{j=1}^{p} \beta_j x_{ij} \ , \quad i = 1,2, \ldots, N, \quad (1)$$

where x_{ij} denotes the value of observation i's jth covariate. It should be noted that $Y \sim N(\mu, \sigma^2)$ and covariates $x_1, x_2, \ldots, x_p$ produce a linear map given by $\eta = \sum_{j=1}^{p} \beta_j x_j$. The link between the systematic and the random components is $\eta = \mu$. Here, the new parameter η and μ are identical and expressed as $\eta_i = g(\mu_i)$ where $g(.)$ is called the link function. Therefore, it is clear that GLM allows two extensions; first, the distribution may come from an exponential family; second, the link function, $g(.)$, may be chosen as any monotonic and differentiable function (Nelder and Baker, 1972).

3.2 Multivariate Adaptive Regression Splines (MARS) Method

For uncovering and complex data patterns, Multivariate Adaptive Regression Splines (MARS) is a popular nonparametric regression method used for the estimation of the general functions of high dimensional arguments. MARS makes no specific assumption about the underlying functional relationship between the response and predictor variables (Barlevy and Fisher, 1993; Frieadman, 1991; Hastie, 1989). The following general model form is represented by a linear combination of basis functions and the intercept as

$$Y = \beta_0 + \sum_{m=1}^{M} \beta_m H_m(\boldsymbol{x}^m) + \epsilon, \qquad (2)$$

where Y is a response variable, ϵ is an error term, which is assumed to have zero mean and a finite variance. Here, β_m are the unknown coefficients for the mth basis function ($m = 1,2,\dots,M$) and for the constant 1 ($m = 0$). The functions H_m ($m = 1,2,\dots,M$) are basis functions taken from a set of M linearly independent basis elements. They can be in a form of main or interaction. For a given data pair $(\boldsymbol{x}_i, y_i)$ $(i = 1,2,\dots N)$, the form of the mth basis function for the multi-predictor case is as follows

$$H_m(\boldsymbol{x}^m) := \prod_{J=1}^{K_m} \left[s_{\kappa_j^m} \cdot \left(x_{\kappa_j^m} - \tau_{\kappa_j^m} \right) \right]_+,$$

where $[q]_+ := \max\{0, q\}$, K_m is the number of truncated linear functions multiplied in the mth basis function, $x_{\kappa\ j}^{m}$ is the input variable corresponding to the jth truncated linear function in the mth basis function, $\tau_{\kappa\ j}^{m}$ is the knot value corresponding to the variable $x_{\kappa\ j}^{m}$, and $s_{\kappa\ j}^{m}$ is the selected sign $+1$ or -1.

4. EMPIRICAL ANALYSIS

4.1 Data Set and Preliminary Analyses

We use monthly data to assess the predictive efficiency of the models and the direction of statistically significant indicators in the US housing market. Based on the guiding literature, we select 11 economic variables (Table 1) and define as substantial factors for the US housing market price variability. Monthly observation of Civilian Unemployment Rate, Consumer Price Index, 30-Year Fixed Rate Mortgage Average in the US, 10-Year Treasury Constant Maturity Rate, US/Euro Foreign Exchange Rate, Effective Federal Funds Rate, Crude Oil Price (West Texas Intermediate), Henry Hub Natural Gas Spot Price, Capacity Utilization (Total Industry), Working-Age Population (Aged 15-64) and Real Disposable Personal Income (Per Capita) are expected to have an impact on the US housing market, which is represented by S\&P/Case-Shiller US National Home Price Index. It is noteworthy to state that inclusion of Capacity Utilization (Total Industry) as exogenous variable and examining also the impact of the recession cycles of the US on its housing market by dummy variables are original in this paper.

Table 1: The variables selected to model the house price index

Variable	Abbreviation
S&P/Case-Shiller U.S. National Home Price	HPI
10-Year Treasury Constant Maturity Rate	TBill
30-Year Fixed Rate Mortgage Average in the United States	FRM
Consumer Price Index for All Urban Consumers: All Items	CPI
US/Euro Foreign Exchange Rate, U.S. Dollars to One Euro	ER
Effective Federal Funds Rate	EFFR
Crude Oil Prices: West Texas Intermediate (WTI)	COP
Henry Hub Natural Gas Spot Price	GSP
Working-Age Population: Aged 15-64: All Persons	WAP
Real Disposable Personal Income: Per Capita	RDI
Capacity Utilization: Total Industry, Percent of Capacity	TCU
Civilian Unemployment Rate	UER
Recession Cycle in the Economy	RC

Since the complexity of housing markets and their specific characteristics that distinguish them from fully competitive markets, it is cumbersome to detect which of these variables may have significant effect on housing markets.

The descriptive statistics given in Table 2 concludes: (i) approximately, half of the variables have right-skewed distributions (TBill, FRM, EFFR, COP, GSP, RDI, UER), (ii) average values of the spot rates range between 1.92% to 3.62% while FRM yields 5.38% over the period January 1999-June 2018, (iii) within the same period, the average of HPI and UER is relatively high, (iv) CPI, COP, WAP and RDI show the highest variability, (v) Jargue-Bera (JB) test, shows none of the variables satisfy the normality assumption.

Table 2: Descriptive statistics of the selected housing market indicators

	HPI	TBill	FRM	CPI	ER	EFFR	COP	GSP	WAP (Mi.*)	RDI (Th.*)	TCU	UER
Mean	151.91	3.62	5.38	210.02	1.21	1.92	59.80	4.58	194.63*	11.58*	77.25	5.95
Median	151.52	3.69	5.40	213.45	1.23	1.15	57.31	4.01	196.97*	1154*	77.33	5.40
St.Dev.	28.03	1.30	1.35	25.36	0.17	2.06	28.01	2.16	9.276*	1.443*	3.16	1.77
Kurtosis	2.44	2.02	2.05	1.72	2.42	2.25	5.11	5.63	2.17	2.11	4.17	2.58
Skewness	-0.32	0.21	0.30	-0.19	-0.29	0.86	0.34	1.50	-0.59	0.04	-0.90	0.90
JB	1	1	1	1	1	1	1	1	1	1	1	1

The dependence between selected variables (Table 3) is measured by Pearson correlation at which the response variable, HPI has the highest association with RDI (79%) and WAP (78%) followed by CPI (74%) and FRM (-56%). Interestingly, although many studies have noted a high and positive relation between house price and unemployment (e.g., Bojan and Darja, 2016; Green and Hendershott, 2001; Oswald, 1999), the data yields meager and negative association (-13%). It is also clear that the interest rates (TBill and EFFR) and FRM are all negatively correlated with HPI. Note that Figure 1 also supports the results that TBill, EFFR, and FRM are negatively correlated with HPI and positively correlated with each other. It is noted that association between TBill and FRM (98%), TBill and EFFR (83%), FRM and EFFR (83%), CPI and RDI (98%) and WAP and CPI (98%) show strong correlation.

The variables chosen for this study are investigated in many empirical studies. An example of high-income increasing CPI is the Lawson boom of the late 1980s, which is followed by the recession of 1981, observed in the United Kingdom (UK). CPI is significantly linked with the TBill, EFFR and FRM. Generally, lower spot rates increase the number of consumers that borrow and increase the consumption in the economy. The consequence is that households have more money to spend; resulting in the growth of the economy and accelerates the CPI. On the contrary, if spot rates increase the households tend to save money and with less disposable income to spend, the economy slows and so inflation decreases. Green and Hendershott (2001) emphasize that older cohorts have both higher homeowner-ship rate and lower unemployment rate than the younger cohorts. Hence, as the population gets older in countries, it is likely to have both higher ownership rates and unemployment rates. This means we may anticipate that the aging communities in countries would generate a negative correlation between homeowner-ship and unemployment. The correlation coefficient between HPI and WAP is also negatively correlated and relatively low.

Table 3: The association among house price index and explanatory variables

	HPI	TBill	FRM	CPI	ER	EFFR	COP	GSP	WAP	RDI	TCU
TBill	-0.52	1									
FRM	-0.56	0.98	1								
CPI	0.74	-0.88	-0.89	1							
ER	0.45	-0.39	-0.40	0.48	1						
EFFR	-0.27	0.83	0.83	-0.67	-0.36	1					
COP	0.41	-0.49	-0.49	0.62	0.82	-0.39	1				
GSP	0.23	0.30	0.31	-0.21	0.35	0.32	0.25	1			
WAP	0.78	-0.88	-0.90	0.98	0.54	-0.69	0.64	-0.13	1		
RDI	0.79	-0.86	-0.87	0.98	0.38	-0.63	0.50	-0.23	0.97	1	
TCU	-0.01	0.40	0.39	-0.27	-0.11	0.65	-0.02	0.28	-0.31	-0.24	1
UER	-0.13	-0.43	-0.44	0.28	0.54	-0.66	0.52	-0.12	0.34	0.16	-0.65

4.2 Parameter Estimation

As none of the variables follow the normal distribution and to reduce the influence of other hidden factors such as auto-correlation and multi-collinearity we normalize the series using

$$X_N = \frac{X - \min(X)}{\max(X) - \min(X)}.$$

We employ Augmented Dickey-Fuller (ADF) test to detect the unit root properties, which indicates that all the series are non-stationary at the level. Therefore, we use the first difference of the series in the analyses.

Case 1: Linear-GLM (L-GLM)

The linear link function yields the estimated model to be

$$\hat{Y}_{HPI} = -0.004 + 0.0788\,CPI + 0.0411\,EFFR - 0.004GSP + 0.070\,WAP + 0.0386\,TCU - 0.0243\,UER + 0.0112\,RC. \quad (3)$$

Here, the model contains only statistically significant variables since we have used stepwise method for modelling. The statistically significant parameter estimates conclude that if an increase occurs in CPI, EFFR, WAP, and RC, there will be an increase in US house prices. On the other hand, a rise in GSP and UER will lead to a decrease in the prices in the US housing market. More specifically, under the assumption of ceteris paribus, we conclude the following:

i. The inflation hedging ability of housing markets is well known. Therefore, in the high inflationary periods house prices increase due to the rise in the levels of housing service in response to community demand (Singell and Lillydahl, 1990). On the other hand, it affects housing markets over the long term. Although most of the people consider the price increase in housing markets as improvement of markets, generally, the reason behind the scene is the inflation. It is because, when we consider the inflation while evaluating the price increase of a house, it will be observed that the real increase will get smaller. Besides, when the inflation rate increases, so do the cost of construction, which causes an increase in house prices. Andrews (2010) gives a proof that inflation has a positive effect on housing markets. However, there exist studies, which claims that inflation is having an adverse impact on house prices. For instance, Follain (1982) investigates the link between inflation and housing markets, and he reports that inflation hurts house prices. In this study, inflation has a positive impact on the US housing market, which conflicts with the results of Follain (1982) but coincides with the findings of Andrews (2010).

ii. Residential investments tended to turn prior to house prices in the business cycle. In recent decades, investments in the housing market have shown high growth in many countries. Low-spot rates have been one of the driving factors as they stimulate the demand in housing markets. Therefore, the lower spot rates cause an increase in house prices and, in turn, stimulate residential investments (Arestis et al., 2009). However, according to the linear GLM, as EFFR increases HPI value increases. Although this result seems to be a contradiction to Arestis et al. (2009), it is economically significant according to two important aspects: First (lending standards), higher spot rates may provide lenders with more of an incentive to make loans and a little bit of a cushion against risk. Second (households' psychology), the expectation of an increase in spot rates will cause a demand increase since people willing to purchase a house before rates go up. Moreover, this situation may increase the quality of mortgages since it pushes people to purchase houses only, they may afford monthly payments.

iii. Prices in housing markets are increased as a result of better employment opportunities and higher incomes enjoyed by residents in an expanding economy since the demand for housing is dependent upon household income. Indeed, the higher economic growth and a rise in the household's income will lead families to spend more on houses. On the other hand, employment and household income are highly dependent on each other (Lerbs, 2011). Thus, housing markets have a healthy relationship with the employment level. Especially, booms in housing investments increase the employment rate, as the construction sector covers more than 20 percent employment gains since early 2000 in the US. GSP is highly related to the energy market and labour market. For instance, some of the population benefited from the construction of energy producing facilities, which are built to allow the export of GSP. The construction sub-sector has led to job opportunities for households. Thus, the job growth has led to an increase in demand for housing. Thus, GSP increases house prices. Equation (3) shows that GSP has a positive effect on the US housing market.

iv. Mankiw and Weil (1989) study on the link between demographics and housing markets, and they conclude that an increase in the new born affects the housing market twenty years later. This means that the working population has a strong relation with housing markets.

v. TCU measures the efficiency of resources by corporations and factories to produce goods in manufacturing, mining, and electricity and gas utilities located in the US. Therefore, capacity utilization highly depends on demand and scheduling production for the most efficient use of facilities in a county. From this point of view, it affects the cost of new houses and the house values in markets. The linear model finds that when the efficiency has an upward trend, the house values will increase. It is an expected result since the increase in TCU increased the income of the consumers and triggered the demand in housing.

vi. The unemployment causes a recession in housing markets, which is also declared in (Schnure, 2005). However, Oswald (1999) and Blanchflower and Oswald (2013) propose that home-ownership increases the unemployment rate as it affects the labour mobility. Contrary to Oswald (1999) and Blanchflower and Oswald (2013), our study reveals that a rise in unemployment causes a decrease in the house prices as in (Branch et al., 2016).

vii. The model also reveals that the US housing market effected by recessions in the economy.

Case 2: Quadratic-GLM (Q-GLM)

The linear relationship moderated by explanatory variables is a simplistic way to explain a dependent variable, but it has a number of drawbacks. For instance, it may not capture certain nonlinear relationships, and it may make no sense for certain parameters. For accurate modelling, the inclusion of the two-way interactions is crucial since mutual influence can be observed in the parameters of the interaction terms. Therefore, as Case 2, we estimate the model for HPI using a quadratic link function that allows interactions among explanatory variables with GLM. Statistically, it yields a polynomial equation that illustrates the influence of statistically significant variables on HPI.

Under the quadratic link function assumption, the estimated model becomes,

$$\hat{Y}_{HPI} = -0.006 + 0.009\,TBill + 0.089\,EFFR + 0.085\,TCU + 0.016\,RC + 0.815\,EFFR \times TCU - 0.0419\,EFFR \times RC - 0.0620\,TCU \times RC - 0.003\,RC^2. \qquad (4)$$

The model illustrated by Equation (4) also contains only statistically significant terms since the stepwise method drop out the non-significant terms. This model involves the main effect of four explanatory variables namely; TBill, EFFR, TCU, and RC, and four interaction terms between variables: EFFR has interactions with TCU and RC; TCU has interactions with

EFFR and RC; RC has interactions with TCU and itself, which are selected according to their relative association to each other. Here, the positive coefficients of the interaction terms suggest that the house prices become more favourable as the variables increases. However, the size and precise nature of these effects are not easy to divine from the examination of the interaction coefficients alone.

The significance of quadratic terms signals that the relationship between HPI and explanatory variables may be non-linear. Therefore, it is cumbersome to interpret the individual coefficients in Case 2, since variables tied to each other. However, intuitively we may interpret the followings from Equation (4):

i. TBill is tied to any of the variables. Thus, its effect may be explained as in the L-GLM case.

ii. However, notice that there are three terms, which essentially contain EFFR. So, if we combine these terms, the aggregate effect of EFFR is being $(0.089 + 0.815 \cdot TCU - 0.0419 \cdot RC)$. Thus, on the contrary to Case 1, for some values of TCU and RC the effect of EFFR is negative. The quadratic model shows that the effect of EFFR depends on the levels of TCU and RC. So, sort of a way the coefficients of TCU and RC adjusting the effective price of EFFR.

iii. The aggregate effect of TCU is determined by $(0.085 + 0.815\, EFFR - 0.0419\, RC)$. The aggregate effect shows that if $EFFR = RC = 0$, the rate of change will be 0.085. The coefficient -0.0419 tells both the direction and steepness of the curvature. Thus, it indicates that RC has a concave down effect on HPI.

iv. Similarly, the aggregate effect of RC has a concave down impact on HPI.

Case 3: MARS

In the construction of the model using MARS method, the maximum number of basis functions ($M_{\max}$) and the highest degree of interactions (K_m) are determined by trial and error. Among many alternative models, the model in Equation (5) is chosen as the best model to fit the US housing market with parameters $M_{max} = 100$ and $K_m = 2$.

$$\hat{Y}_{HPI} = 0.0069 - 0.2387\, H_1 - 0.2369\, H_2 - 0.1202\, H_3 + 0.3872\, H_4 + 0.0640 H_5 - 0.1546 H_6 - 0.0161 H_7 + 0.0083 H_8 - 0.0749\, H_9 + 3.9605 H_{10} - 2.3779 H_{11} - 1.2907 H_{12} + 2.0095 H_{13} - 2.0001\, H_{14} + 2.3246\, H_{15} - 19.5806 H_{16} + 7.3737 H_{17} - 0.3601 H_{18} + 0.3043 H_{19} - 0.9922 H_{20} + 0.0625 H_{21} + 0.1166\, H_{22} + 0.4982\, H_{23} - 9.3394 H_{24} + 20.0889 H_{25} - 10.6264 H_{26} + 0.0320 H_{27} - 0.1394 H_{28}. \qquad (5)$$

MARS model (Equation (5)) includes a total of 28 basis functions to explain the inherently complex nature of the US housing market data. It is because MARS models are developed automatically and adaptively requiring less applicant expertise. MARS method produces the powerful prediction by building models by over all possible combinations of explanatory variables and all values of each variable as candidates of knots automatically. Therefore, there is a large group of knot points in Case 3.

Besides its complexity, the MARS model is capable of exploring both linear and nonlinear relationships between variables through the additive and interaction basis functions determined as above. The most frequently used variables in the MARS model are GSP, COP, RC, ER, and EFFR. To assess the relative importance of independent variables, the complete MARS model is evaluated in detail concerning both additive basis functions such as the first eight basis functions (H_1 to H_8) and interaction basis functions (interaction between only two independent variables) such as H_9 to H_{28}. It should be noted that the knot values of basis functions are the first difference of the series. The basis functions, H_m, $m = 1,2,\cdots,28$, in Equation (5) are as follows:

$H_1 = \max\{0, ER - 0.06\}$,

$H_2 = \max\{0, EFFR + 0.02\}$,

$H_3 = \max\{0, 0.01 - EFFR\}$,

$H_4 = \max\{0, EFFR - 0.01\}$,

$H_5 = \max\{0, COP + 0.07\}$,

$H_6 = \max\{0, 0.03 - GSP\}$,

$H_7 = \max\{0, 1 - RC\}$,

$H_8 = \max\{0, RC - 1\}$,

$H_9 = ER \cdot \max\{0, RC - 1\}$,

$H_{10} = \max\{0, CPI - 0\} \cdot \max\{0, 0.04\, COP\}$,

$H_{11} = \max\{0, 0.06 - ER\} \cdot \max\{0, COP - 0.04\}$,

$H_{12} = \max\{0, ER + 0.02\} \cdot \max\{0, GSP - 0.03\}$,

$H_{13} = \max\{0, ER + 0.01\} \cdot \max\{0, GSP - 0.03\}$,

$H_{14} = \max\{0, ER - 0.02\} \cdot \max\{0, GSP - 0.03\}$,

$H_{\{15\}} = \max\{0, ER - 0.04\} \cdot \max\{0, GSP - 0.03\}$,

$H_{16} = \max\{0, 0.01 - EFFR\} \cdot \max\{0, COP - 0.03\}$,

$H_{17} = \max\{0, 0.01 - EFFR\} \cdot \max\{0, COP - 0.01\}$,

$H_{18} = \max\{0, 0.01 - EFFR\} \cdot \max\{0, RC - 1\}$,

$H_{19} = \max\{0, 0.04 - COP\} \cdot \max\{0, 0.35 - GSP\}$,

$H_{20} = \max\{0, 0.04 - COP\} \cdot \max\{0, UER - 0.02\}$,

$H_{21} = \max\{0, 0 - COP\} \cdot \max\{0, 1 - RC\}$,

$H_{22} = \max\{0, COP - 0\} \cdot \max\{0, 1 - RC\}$,

$H_{23} = \max\{0, GSP - 0.03\} \cdot \max\{0, RDI - 0\}$,

$H_{24} = \max\{0, GSP - 0.09\} \cdot \max\{0, 0.03 - TCU\}$,

$H_{25} = \max\{0, GSP - 0.1\} \cdot \max\{0, 0.03 - TCU\}$,

$H_{26} = \max\{0, GSP - 0.11\} \cdot \max\{0, 0.03 - TCU\}$,

$H_{27} = \max\{0, 0.27 - GSP\} \cdot \max\{0, 1 - RC\}$,

$H_{28} = \max\{0, 0 - TCU\} \cdot \max\{0, 1 - RC\}$.

The results of the MARS model indicate that independent variables ER, COP, GSP, RC and EFFR involved in both types of basis functions have the highest effect on the dependent variable (HPI) when compared with the other independent variables. The knot point for basis function H_3 is 0.01. The interpretation of this basis function is that as EFFR values get smaller values than 0.01, HPI decreases. On the other hand, the basis function H_{16} contains the basis function H_3 to express the interaction between the independent variables EFFR and COP. Similarly, the basis function H_{18} represents the interaction between the independent variables EFFR and RC.

Among basis functions H_{16}, H_{17}, H_{24}, H_{25}, and H_{26} have the largest effect on the US housing market (HPI). Here, only basis functions $H_{17} = \max\{0, 0.01 - EFFR\} \cdot \max\{0, COP - 0.01\}$ and $H_{25} = \max\{0, GSP - 0.1\} \cdot \max\{0, 0.03 - TCU\}$ have a positive impact on the US housing market. While the basis function H_{17} contains the interaction between EFFR and COP, the basis function H_{25} has the interaction between GSP and TCU.

The MARS model includes 28 terms while Q-GLM and L-GLM include much less terms. However, there exist some limitations on using basis functions due to their interaction forms. For instance, basis functions H_{12}, H_{13}, H_{14} and H_{15} contain the main function, which is $\max\{0, GSP - 0.03\}$. Hence, some of them have no effect on HPI since they get zero values related to ER's value. On the other hand, basis functions H_{24}, H_{25} and H_{26} includes the main function, which is $\max\{0, 0.03 - TCU\}$. Therefore, their values depend on the change on GSP values. For instance, while GSP value is less than 0.09 they have no effect on HPI, they all have an effect on HPI when GSP value greater than 0.11. On the other hand, if GSP value is between 0.1 and 0.11 only H_{25} and H_{24} have effect on HPI. Similar cases observed for basis functions H_{16}, H_{17} and H_{21}, H_{22}. In this respect, some of the basis functions do not affect HPI at the same time.

Figure 3 is introduced to visualize the monthly observed and predicted HPI values for the period 1999-2018. The data contains the latest global financial crisis. The MARS, Q-GLM and L-GLM captures/detects and quantifies the crisis since they include Recession Cycles (RC). We observe that there is compliance with the observed and predicted HPI values even during the global financial crisis for all models. The evolution of the anticipated prices indicates that all models' predictions are relatively significant and they can be used to determine the direction of the prices for the US housing market.

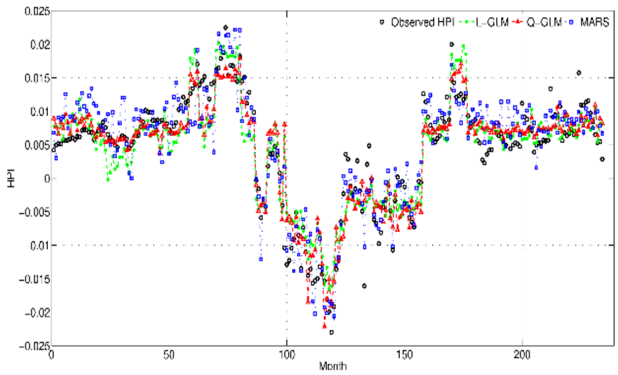

Figure 3: MARS and GLM model fits on real data (1999-2018)

4.3 Performance of the models

To evaluate and compare the performances of models (MARS, L-GLM, Q-GLM), we divide the data set into two parts: we use 175 observations as a training sample and 59 observations as a validation sample. The first part of these samples is used for the estimation of the model parameters, and the rest is employed in the validation of the models. The prediction results from MARS and GLMs are further evaluated concerning well-known performance measures such as Mean Absolute Error (MAE), Root Mean Square Error (RMSE), Coefficient of Determination (R^2), and Proportion of Residuals within Three Sigma (PWI). The smaller values for MAE and RMSE indicate a reasonable estimation of the model parameter. There are no well-defined bounds on these values of the performance measures. On the other hand, the other measures (R^2, PWI) indicate a better performance when their values are closer to 1.

We summarize the models' performance measures and their forecasting power for comparing the models' efficiency in Table 4. The table reveals that for the training dataset, the Q-GLM shows the best performance according to most of the measures (MAE, RMSE and R^2). However, there is no significant difference among the performance of the models for training data. We also see that MARS model performs much better than Q-GLM and L-GLM according to almost all measures for the validation dataset. From these results, we may conclude that, in both the training sample and the validation sample, the MARS model has a good prediction capability and discovers the main structure of the data very well. It is because the MARS model uses the power of piecewise functions in capturing the data structure. As a result of this, MARS can be successfully applied in the validation sample after the model building procedure.

Table 4: Performances of MARS and GLMs for on train and test data sets

	Train Set			Validation set		
Performance measure	MARS	Q-GLM	L-GLM	MARS	Q-CLM	L-GLM
MAE	0.0036	0.0031	0.0035	0.0038	0.0036	0.0042
RMSE	0.0044	0.0042	0.0047	0.0051	0.0053	0.0057
R^2	0.76	0.79	0.74	0.66	0.63	0.57
PWI	0.97	0.96	0.95	0.96	0.93	0.97

5. SUMMARY AND CONCLUSION

This study provides an insight into factors connected with housing markets. It has two main contributions to the literature. The first is to highlight the link between the critical economic indicators and housing markets with their direction of the interaction to explain the price fluctuations in housing markets. That is, housing markets have a strong relationship with other markets, and its potential volatility may have a dramatic impact on overall economies even in some cases it may lead to global crises. The second contribution is to shed new light on the mechanism of housing markets and to employ statistical models that can be used to predict housing markets behaviour. The empirical models provide a coherent set of empirical and prediction results. The models also confirm the importance of changes in national economic conditions, unemployment, working-age population, inflation, income, interest rates, gas spot price, crude oil prices, capacity utilization, recessions in economies affect housing markets.

In addition to these contributions, it is worth to emphasize that it is possible to forecast the direction of changes in housing prices by using the models within the current study. The models and analysis of a potential influence of particular macro-economic indicators on housing markets are driven by using the GLMs and MARS. The explanatory variable selection and modelling methods are the keys for constructing such models that may show variation due to the characteristics of countries. Thus, concerning explanatory variables included in this study, many other explanatory variables may be added based on the country-specific characteristics.

The modelling methods presented in the study, make the use of the GLM and MARS for the first time in housing markets under such broad exogenous factors. The methods suggest a significant advance in nonparametric modelling with macro-economic indicators for the US housing market. In the usual multi-regression modelling and GLM, for extensive observations and a large number of predictor variables, the search can be computationally time-consuming. However, the MARS algorithm automatically achieves the selecting predictors/knots in an efficient, and adaptive way and finds the most straightforward and best model that balances the over-fitting and under-fitting for the model. Therefore, MARS is superior to GLMs from the prediction performance point of view and saving significant amounts of computational time.

REFERENCES

Adams, Z. and Füss, R. (2010). Macroeconomic determinants of international housing markets. Journal of Housing Economics, 19(1):38–50.

Agnello, L. and Schuknecht, L. (2011). Booms and busts in housing markets: determinants and implications. Journal of Housing Economics, 20(3):171–190.

Andrews, D. (2010). Real house prices in oecd countries: The role of demand shocks and structural and policy factors. OECD Economic Department Working Papers, 1(831):0–34

Arestis, P., Mooslechner, P., and Wagner, K. (2009). Housing Market Challenges in Europe and the United States. Palgrave Macmillan UK.

Aspachs-Bracons, O. and Rabanal, P. (2010). The drivers of housing cycles in Spain. SERIEs, 1(1):101–130.

Barlevy, G. and Fisher, J. (1993). Fast mars, technical report 110. Technical report, Stanford Linear Accelerator Center.

Blanchflower, D. G. and Oswald, A. J. (2013). Does high home-ownership impair the labor market? Technical report, National Bureau of Economic Research.

Bojan, G. and Darja, K. G. (2016). Influence of macroeconomic factors on prices of real estate in various cultural environments: Case of Slovenia, Greece, France, Poland and Norway. Procedia Economics and Finance, 39:597 – 604.

3rd Global Conference on Business, Economics, Management and Tourism.

Branch, W. A., Petrosky-Nadeau, N., and Rocheteau, G. (2016). Financial frictions, the housing market, and unemployment. Journal of Economic Theory, 164:101–135.

Carson, R. T. and Dastrup, S. R. (2013). After the fall: An ex post characterization of housing price declines across metropolitan areas. Contemporary Economic Policy, 31(1):22–43.

Chauvet, M., Gabriel, S., and Lutz, C. (2016). Mortgage default risk: New evidence from internet search queries. Journal of Urban Economics, 96:91–111.

Diewert, W. E., Nakamura, A. O., and Nakamura, L. I. (2009). The housing bubble and a new approach to accounting for housing in a cpi. Journal of Housing Economics, 18(3):156–171.

Dokko, J. K., Keys, B. J., and Relihan, L. E. (2015). Affordability, financial innovation, and the start of the housing boom. Working paper.

Englund, P. and Ioannides, Y. M. (1997). House price dynamics: an international empirical perspective. Journal of Housing Economics, 6(2):119–136.

Follain, J. R. (1982). Does inflation affect real behavior: The case of housing. Southern Economic Journal, 48(3):570–582.

Friedman, J. H. (1991). Multivariate adaptive regression splines. The Annals of Statistics, 19(1):1–67.

Gadi Barlevy, J. D. M. F. (2011). Mortgage choices and housing speculation. Available at SSRN, 5(1):1–63.

Glaeser, E. L. and Parker, J. A. (2000). Comments and discussion. Brookings Papers on Economic Activity, 2000(2):146–159.

Green, R. K. and Hendershott, P. H. (2001). Home-ownership and unemployment in the US. Urban Studies, 38(9):1509–1520.

Guirguis, H. S., Giannikos, C. I., and Anderson, R. I. (2005). The us housing market: Asset pricing forecasts using time varying coefficients. The Journal of Real Estate Finance and Economics, 30(1):33–53.

Hastie, T. (1989). Flexible parsimonious smoothing and additive modeling: Discussion. Technometrics, 31(1):23–29.

Iacoviello, M. (2005). House prices, borrowing constraints, and monetary policy in the business cycle. American Economic Review, 95(3):739–764.

Johnes, G. and Hyclak, T. (1999). House prices and regional labor markets. The Annals of Regional Science, 33(1):33–49.

Lambertini, L., Mendicino, C., and Punzi, M. T. (2017). Expectations-driven cycles in the housing market. Economic Modelling, 60:297–312.

Leamer, E. (2013). Chapter 41 - housing is the business cycle. In Caprio, G., Beck, T., Claessens, S., and Schmukler, S. L., editors, The Evidence and Impact of Financial Globalization, pages 589 – 643. Academic Press, San Diego.

Lerbs, O. (2011). Is there a link between homeownership and unemployment?: Evidence from German regional data. International Economics and Economic Policy, 8(4):407–426.

Malpezzi, S. (1999). A simple error correction model of house prices. Journal of Housing Economics, 8(1):27–62.

Mankiw, N. G. and Weil, D. N. (1989). The baby boom, the baby bust, and the housing market. Regional Science and Urban Economics, 19(2):235–258.

Nelder, J. A. and Baker, R. J. (1972). Generalized Linear Models. Wiley Online Library.

Ortalo-Magne, F. and Rady, S. (2006). Housing market dynamics: On the contribution of income shocks and credit constraints. The Review of Economic Studies, 73(2):459–485.

Oswald, A. J. (1999). The housing market and Europe's unemployment: A non-technical paper. Technical report, Department of Economics University of Warwick.

Poterba, J. (1991). House price dynamics: The role of tax policy. Brookings Papers on Economic Activity, 22(2):143–204.

Rapach, D. E. and Strauss, J. K. (2009). Differences in housing price forecastability across US States. International Journal of Forecasting, 25(2):351 – 372.

Schnure, C. (2005). Boom-bust cycles in housing: The changing role of financial structure. IMF Working Paper, 200(5):1–27.

Singell, L. D. and Lillydahl, J. H. (1990). An empirical examination of the effect of impact fees on the housing market. Land Economics, 66(1):82–92.

Valentini, P., Ippoliti, L., and Fontanella, L. (2013). Modeling US housing prices by spatial dynamic structural equation models. The Annals of Applied Statistics, 7(2):763–798.