STAT 33700 Multivariate Time Series Analysis

Final Project

Analysis & Prediction of Chicago's Housing Market

Zhi Rong Tan

1. Introduction

Housing costs are an immediate concern for many American families and policymakers. High prices may have an adverse impact on cost of living, while low prices can be problematic for long-term homeowners who rely on housing wealth to supplement retirement savings. Since property market stability is closely linked to macroeconomic and financial stability, it is crucial to understand housing trends and variables affecting its movement.

Fundamentally, housing markets are governed by the laws of demand and supply. According to Glaeser and Gyourko (2008), the strong negative correlation between housing permits (an indicator of housing supply) and the level of housing prices across markets reveals the importance of supply-side conditions. On the other hand, Mankiw and Weil (1989) studied the effect of demographic changes on housing prices, and suggested that the decline in the fraction of the American population in the prime house buying age bracket during the 1990s had caused a significant fall in real housing prices.

Similarly, Johnes and Hyclak (1996) explored the impact that the housing and labor markets have on each other in 4 US cities – Houston, Milwaukee, Fort Lauderdale, and Hartford, and discovered that the long-term determinants of house prices are nominal gross city product and the labor force size. Evidence demonstrated that in all 4 cities, unemployment and labor force changes affect housing prices, while the converse is also true – house prices have an effect on the labor force size. Johnes and Hyclak's research also highlighted that housing supply increases, as measured by building permits, had a depressing effect on housing prices, but house prices were interest inelastic in both the long and short run.

Given that this research was performed in the 1990s and before the Great Financial Crisis, an interesting extension is to understand how the relationship between labor force and housing supply with housing prices has changed during the recovery period post-recession, and to supplement the research with new cities, such as Chicago.

1.1. Chicago's Housing Trends

The real estate market in Chicago – the 3rd largest US metropolitan area, has been the one of the slowest to recover since the Great Recession housing bubble burst. Chicago home prices were 19% below their pre-crash levels in 2017, and are not expected to hit peak values until 2020 (Murray & Schuetz, 2018). Although there has been a steady increase in local jobs, the Chicago metro area was the only city in the 10 largest metro areas to experience population decline in 2018 due to persistent out-migration. More concerning is the fact that the population decline is driven by working-age Illinoisans, the population group responsible for property demand. Furthermore, Illinois homeowners are subjected to the second highest property tax in the states, and are currently enduring the largest permanent income tax hike in Illinois history.

These trends, if they are to continue, will lead to a slowdown in Chicago's housing market compared to the rest of the States. Beyond these insights, this paper shall explore other current factors affecting Chicago's housing market, and predict its future trajectory.

1.2. Housing Trends across Cities

The rate of housing market growth, and its relationship with other city-level macroeconomic indicators differ from region to region. Capozza et. al. (2002) found that the extent of momentum and reversion of housing prices in US metropolitan areas varies with location, for example, between coastal and inland cities. Divounguy (2018) also discussed how metropolitan areas around the Great Lakes have lower neighborhood price-income ratios than either the coastal or inland regions – the West Coast and northeastern regions have one of the highest housing prices, while the Midwest and Southern regions are able to maintain a stable and affordable housing market, due to differing conditions and policies implemented at the city-level. More specifically, in San Francisco, high housing demand due to a booming job market, coupled with zoning regulations limiting supply have led to high housing prices – in recent years, the Bay Area added 480 000 private-sector jobs but only around 50 000 housing units.

Nonetheless, despite varying conditions across cities, there are also common factors that drive price growth in all cities, such as recovery from the recession, and other federal policies. Can knowledge of one city's housing prices improve understanding and prediction of the housing prices in another city? This paper will observe how similar or different housing markets for America's largest metropolitan areas are, and the performance of metro-level housing markets in the aftermath of the recession.

2. Objectives and Methodology

The paper has 2 objectives. First, we will use time series techniques to analyze how key factors such as housing supply and labor force affect housing prices at the city level and dictate the future movement of Chicago's housing market.

In the selection of factors to analyze, we explored macroeconomic factors in our informal work, such as *long-term and short-term interest rate*, *industry productivity growth*, and *Economic Conditions Index* (which measures economic growth at the city level), but did not include them in this paper since they do not lead to significant results.

Fundamental factors influencing housing prices are the Chicago metropolitan labor force and housing supply. For city-level housing supply, the closest proxy we can find for Chicago that is sufficiently discrete is the *New Private Housing Units Authorized by Building Permits*, but informal work failed to produce a reasonable time series model. Due to the perceived importance of housing supply as a factor, we will proceed with utilizing the *National Housing Supply* Time series data to investigate the effects of supply on Chicago's housing prices. Note that a lack of significant relationship between national housing supply and Chicago's housing prices may be due to 2 issues – either the national housing supply is a poor proxy for Chicago's housing supply, or the implausibility of housing supply as a factor affecting housing prices. To provide greater clarity if that occurs, we include the National housing prices in our multivariate time series. Including national housing prices also enable us to determine the relationship between Chicago's and the national housing prices.

During our research, a tri-variate time series dataset consisting of Chicago's Housing Price, Chicago's Labor Force, and National Housing Supply did not produce a stable time series model. Hence, we split the analysis into 2 time series models: the first a bivariate data consisting of **[1: Chicago's Housing Price, 2: Chicago's Labor Force]**, and the other one consisting of **[1: National Housing Price, 2: Chicago's Housing Price, 3: National Housing Supply]**.

Second, we will compare Chicago's housing prices with other cities' housing prices. Due to the range of available data, we only consider major US metropolitan areas, and choose 2 cities – San Francisco, which has one of the hottest property markets in the US, and Detroit – a city in the same Midwestern region as Chicago that was badly hit by the housing crisis in 2007, but has healthy recovery in the recent few years. By comparing the price index across the 3 cities, we can determine whether changes in housing prices are driven primarily by a common underlying trend, or by factors that are specific to certain regions.

For both portions, we will check for co-integration, fit a *VAR(p)* model, and calculate the impulse response function to determine how a component of the data will react to an exogenous shock in another component over time. Thereafter, we shall test for granger causality between components to better appreciate the relationship between our variables.

Finally, we hope to obtain an accurate forecast of Chicago's housing prices using 2 methods:

1. Time series model of Chicago housing prices with supply and labor force
2. Time series model of Chicago, San Francisco, Detroit's housing prices

The combination of both models would paint a clearer picture of future movement in Chicago's housing market.

3. Data

We will perform 3 separate multivariate time series model, consisting of 6 different time series data. All the data are obtained from the Federal Reserve Bank of St. Louis Economic Data.

Acknowledging that seasonality exists in housing data, we pick datasets that have been seasonally adjusted. All 6 time series data are measured quarterly, starting from **2000 Q1** till **2018 Q3**. The range includes data from 8 years before, to 10 years after the Global Financial Crisis of 2008.

The 6 univariate time series are:

1. S&P/Case-Shiller IL-Chicago Home Price Index
2. S&P/Case-Shiller U.S. National Home Price Index
3. Monthly Supply of Houses in the United States
4. Civilian Labor Force in Chicago-Naperville-Elgin Metropolitan Area
5. S&P/Case-Shiller CA-San Francisco Home Price Index
6. S&P/Case-Shiller MI-Detroit Home Price Index

For (1), (2), (5), (6), the housing prices for each city/nationally are measured with a housing index, normalized to January 2000 = 100. The S&P CoreLogic Case-Shiller Home Price Indices track the price path of a typical *single-family home* pairs for thousands of individual homes within the given region, and are calculated using a 3-month moving average. Percentage change in the index reflects percentage change in housing market prices given a constant level of quality and house size. For example, an index value of 150 implies a 50% appreciation rate since January 2000 for a constant home in the target region. Since the prices are all normalized to 100 in 2000 Q1, we are only comparing the growth of prices since 2000, and not of the absolute price.

For (3), the Monthly Supply for Houses in USA is given by the U.S. Census Bureau, and measures the ratio of houses for sale (in the for-sale inventory) to houses sold. It indicates how long the current for-sale inventory would last given the existing sales rate if no additional houses were built.

For (4), the Civilian Labor Force in Chicago Metropolitan Area is given by the US Bureau of Labor Statistics, and measures the labor force size in units of people, seasonally adjusted.

For all the time series data, I take the logarithm to keep the variability stable.

a) In the first model, we analyze the impact of (national) supply on Chicago housing prices using the multivariate data combining (1), (2), (3).
b) In the second model, we analyze the impact of Chicago labor force on Chicago housing prices using the bivariate data combining (1) and (4).
c) In the final model, we compare the housing prices across different cities using the multivariate data combining (1), (5), (6).

4. Analysis

4.1. Model 1 – Impact of Housing Supply

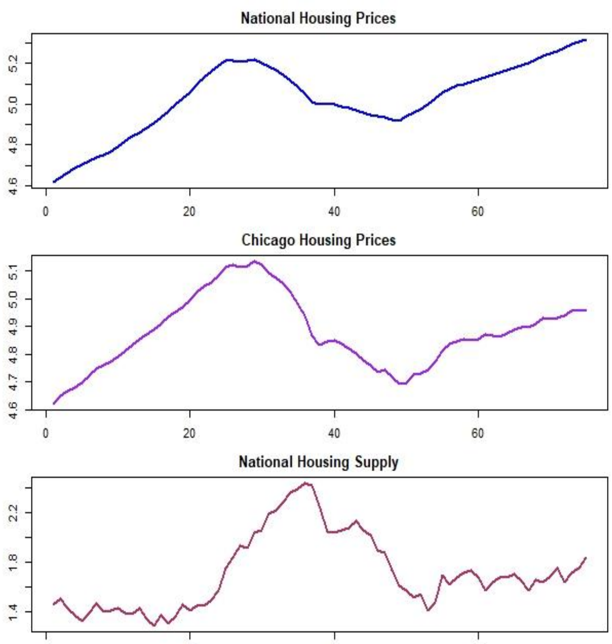


Figure 1: Time Series Plot of logged Data

To begin, let us plot the logged tri-variate time series data of National housing prices (*National*), Chicago housing prices (in price index) (*Chicago*), and national housing supply (*Supply*).

Then, we check for the possibility of unit roots for each component. First, I estimated the AR order for each univariate time series, before performing the Augmented Dickey-Fuller test, obtaining a p-value of 0.81, 0.78, 0.67 for *National*, *Chicago* and *Supply* respectively. This implies that unit root exists for all 3 components, since we did not reject the null hypothesis.

Using the optimal VAR order of 2 by AIC and BIC, we performed the Johansen test for co-integration (including a constant term), and did not reject the null hypothesis that r ≤ 1, with a test statistic of 14.42 against the critical value of 15.67 at $\alpha = 0.05$. We reject the null hypothesis that r = 0 even at 1% significance level.

Due to co-integration, we proceeded with an error correction model below.

$$\boldsymbol{w_t} = (1 \quad -0.67 \quad -6.06)\, \boldsymbol{z_{t-1}}$$
$$\Delta \boldsymbol{z_t} = \boldsymbol{\alpha}\, \boldsymbol{w_t} + \boldsymbol{\phi_0} + \boldsymbol{\phi_1} \Delta \boldsymbol{z_{t-1}} + \boldsymbol{a_t}$$

$$\boldsymbol{\alpha} = (\,0.0046 \quad 0.0021 \quad -0.017\,)^T \qquad\qquad \boldsymbol{\phi_0} = (\,0.088 \quad 0.035 \quad -0.317\,)^T$$

$$\boldsymbol{\phi_1} = \begin{pmatrix} 0.81 & -0.32 & -0.002 \\ 1.04 & -0.19 & -0.015 \\ 1.67 & 0.84 & 0.056 \end{pmatrix} \qquad\qquad \boldsymbol{\Sigma_a} = 10^{-4}\begin{pmatrix} 0.285 & 0.220 & -1.04 \\ 0.220 & 0.950 & 0.047 \\ -1.04 & 0.047 & 55.6 \end{pmatrix}$$

The residual plots, cross-correlation matrix results, Ljung-Box statistics show that the error correction model is sufficient at 5% significance level.

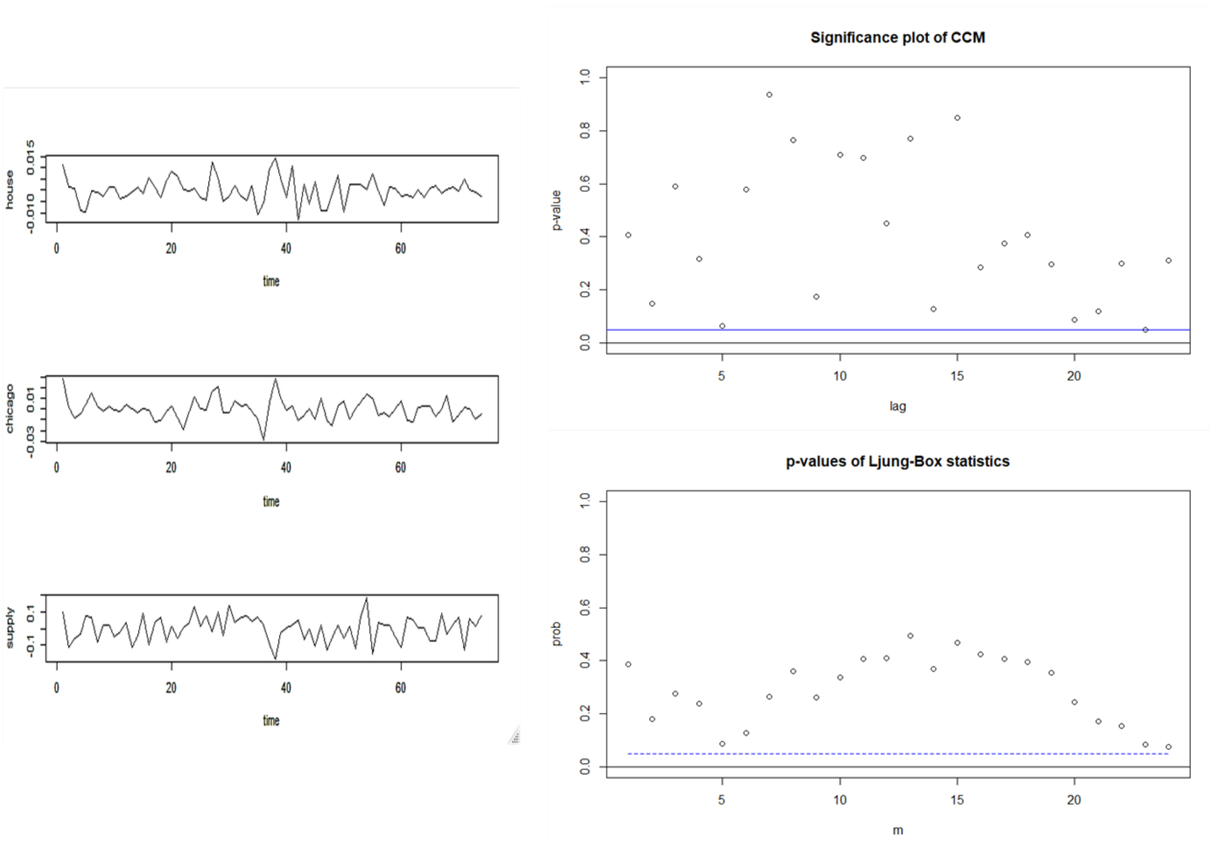


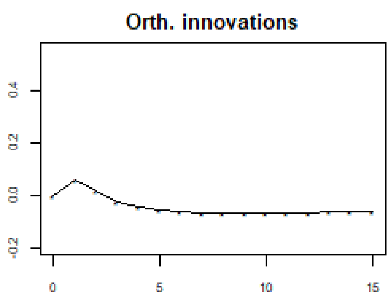
Figure 2: Model Checking for Error-Correction Model

In order to understand how the changes in one variable affect the others marginally, we perform an Impulse Response Function analysis by first transforming the ECM model into its level representation. Note that since the time series is not stationary, the MA representation may not exist, but we can compute the ψ-weights recursively to determine the IRF. The non-stationarity of Z_t also implies that the IRF may not decay to 0.

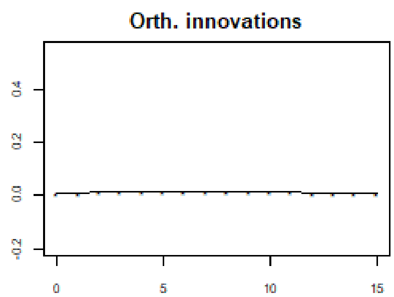
Since our analysis focuses on understanding how supply affects Chicago's housing market, we will mainly look at the IRF (orthogonal innovation) of supply on Chicago's housing prices. We also analyze the relationship between national and Chicago's housing prices. The plots are shown on the next page.

Given a positive shock in national housing supply, observe that Chicago prices responded positively for the first 3 periods, before the instantaneous impact steadily decreases to slightly below 0. The initial positive response may be explained perhaps by market optimism and perceived economic boom driving growth in prices after an initial boom in supply, while the eventual drop below 0 is likely housing prices responding to excessive supply in the longer term. On the other hand, supply do not respond to a shock in Chicago's housing prices at all.

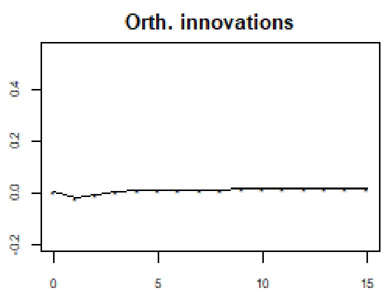
Also notice that national housing prices react positively to shocks in Chicago's prices, but the IRF of Chicago to National is slightly negative at the beginning, before becoming almost flat. This may be because Chicago's housing prices is a part of the national housing price index, but positive shocks to national prices may not directly positively affect Chicago's prices. Indeed, recall that Chicago is noted to have one of the slowest recoveries since the recession, and changes in *Chicago* may be the reverse of what is seen in other cities, as captured by the national housing price index – growth in the national housing prices may not directly lead to growth in Chicago's prices.



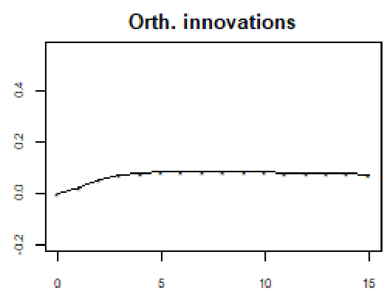
Chicago to shock in Supply



Supply to shock in Chicago



Chicago to shock in National



National to shock in Chicago

We also verify the relationship through a granger causality test of each component. A time series X is said to granger-cause Y if values of X provide statistically significant information about future values of Y.

Hypothesis	P-Value	Conclusion
Supply granger-caused Chicago	0.009	Yes
National granger-caused Chicago	0	Yes
Chicago granger-caused Supply	0.16	**No**
Chicago granger-caused National	0.0023	Yes
Supply granger-caused National	0.0027	Yes

From the test, we see that the granger causality between National housing prices and Chicago housing prices is bi-directional, while supply also granger-caused both national and Chicago housing prices, validating the trends in the IRF and literature.

Finally, we perform a forecast of how the Chicago's housing prices and the national supply will proceed after 2018 Q3, for the next 3 years (12 quarters).

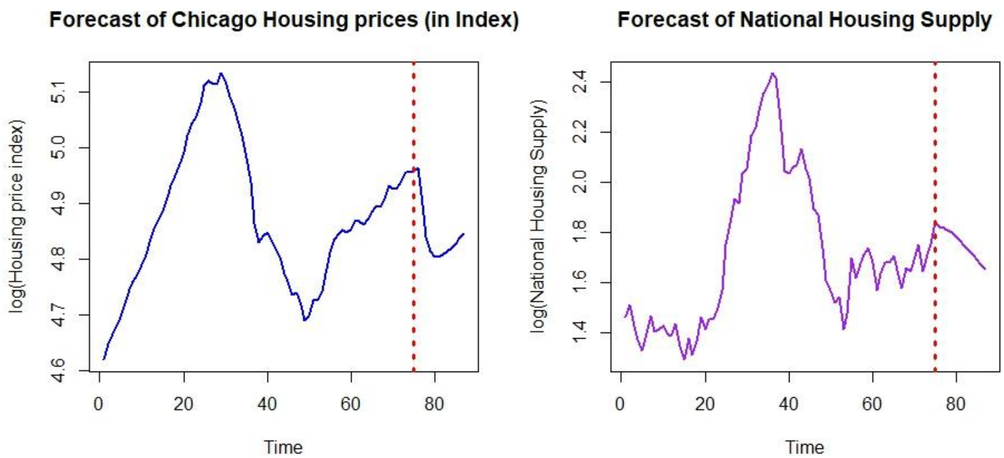


Figure 3: Forecast of Chicago Housing Price Index, National Housing Supply (Model 1)

Observe that the logged national housing supply is forecasted to decrease steadily after 2018 Q3. Similarly, the logged Chicago housing price index is forecasted to increase slightly for the first quarter, before decreasing, then steadily rising again. We will compare this model's forecast with other models in *section 4.4*.

4.2. Model 2 – Impact of Local Labor Force

Now, we will analyze the impact of Chicago's local labor force (Labor) on Chicago's housing price index (Chicago). To begin, let us plot both the logged time series data and first order differenced time series of Chicago's labor force (units of people), shown in *Figure 4*. The time series plot for Chicago housing prices is shown in other sections.

We proceed like in the previous section, checking for the possibility of unit root. After obtaining the AR order for each univariate time series, I performed the Augmented Dickey-Fuller test, obtaining a p-value of 0.78 and 0.70 for *Chicago* and *Labor* respectively, implying that unit root exists for both components.

We then found the optimal VAR order of 4 by BIC, and performed the Johansen test for co-integration (including a constant term). However, with a test statistic of 17.48 against a 10% significance level critical value of 17.85, we failed to reject the null hypothesis that r = 0, i.e. there is no co-integration even at 10% significance level, suggesting a lack of common trend between Chicago housing price index and Chicago labor force. In order to eliminate the issue of unit root, we continue by differencing the multivariate time series. The plot of the new series is shown in *Figure 4*.

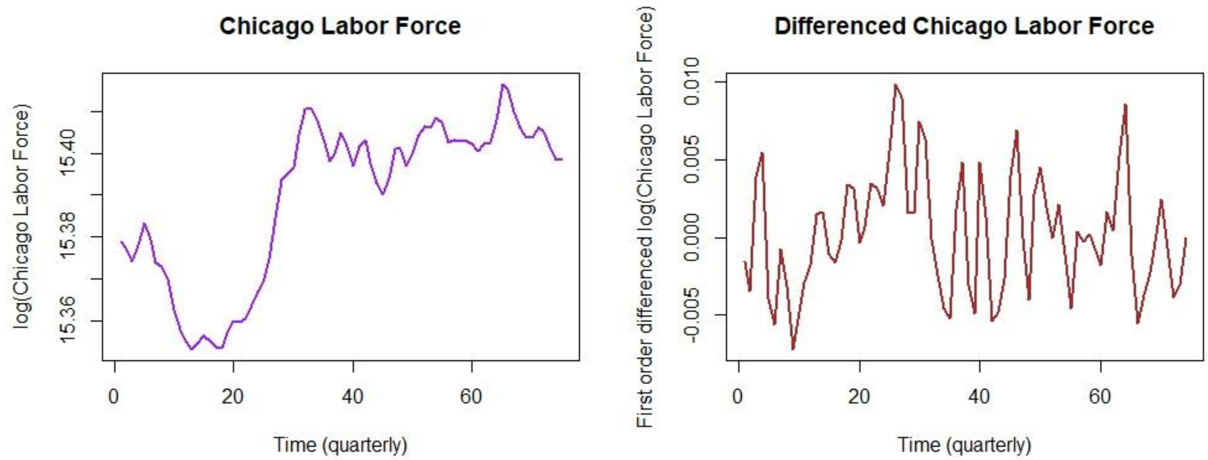


Figure 4: Time Series Plot of logged Chicago Labor Force

Thereafter, we check to ensure the differenced time series does not contain any unit roots. Indeed, the Augmented Dickey-Fuller test for the differenced time series returns a p-value of 0.041 and 0.0105 for Chicago and Labor respectively, implying we can reject the null hypothesis that unit root exists.

Before proceeding, we shall check for Granger Causality between the 2 components. We selected a VAR order of 3 using the BIC and HQ criterion for the Granger causality test.

Hypothesis	P-Value	Conclusion
Chicago granger-caused Labor	0.517	**No**
Labor granger-caused Chicago	0.718	**No**

Interestingly, there is no granger causality in either direction, and hence, knowledge of Chicago's labor force may not improve the accuracy of forecast for Chicago's housing prices, and vice versa.

We fitted a VAR(3) model to verify this, and obtained a model with an AIC of –20.195 and BIC of –19.82. In order to simplify the model and remove parameters that are not statistically significant, we refined the model with a threshold of 1.64, and obtained a new model with an AIC of –20.29 and BIC of –20.11. We will continue with the refined model, shown below, since it has a lower BIC and AIC. Note that the first component is differenced Chicago housing price index, while the second component is the differenced Chicago labor force.

$$\Delta \boldsymbol{z_t} = \boldsymbol{\phi_0} + \boldsymbol{\phi_1} \Delta \boldsymbol{z_{t-1}} + \boldsymbol{\phi_2} \Delta \boldsymbol{z_{t-2}} + \boldsymbol{\phi_3} \Delta \boldsymbol{z_{t-3}} + \boldsymbol{a_t}$$

$$\boldsymbol{\phi_0} = \begin{pmatrix} 0 & 0 \end{pmatrix} \qquad \boldsymbol{\phi_1} = \begin{pmatrix} 0.814 & 0 \\ 0 & 0.855 \end{pmatrix} \qquad \boldsymbol{\phi_2} = \begin{pmatrix} -0.368 & 0 \\ 0 & -0.697 \end{pmatrix}$$

$$\boldsymbol{\phi_3} = \begin{pmatrix} 0.38 & 0 \\ 0 & 0.105 \end{pmatrix} \qquad \boldsymbol{\Sigma_a} = 10^{-6}\begin{pmatrix} 176 & 4.22 \\ -4.22 & 7.49 \end{pmatrix}$$

From the model parameters above, we note that for all ϕ_k , $k = 1,2,3$, $\phi_{k,(1,2)}$ and $\phi_{k,(2,1)} = 0$. Hence, this supports our previous discovery that there is no granger causality between Chicago housing price index and Chicago labor force. Changes in Chicago housing prices are not affected by changes in Chicago labor force during previous time periods, and vice versa.

Residual Analysis (results shown on the next page) via the residual plots, cross-correlation matrix results and Ljung-Box statistics suggest that the refined model is adequate.

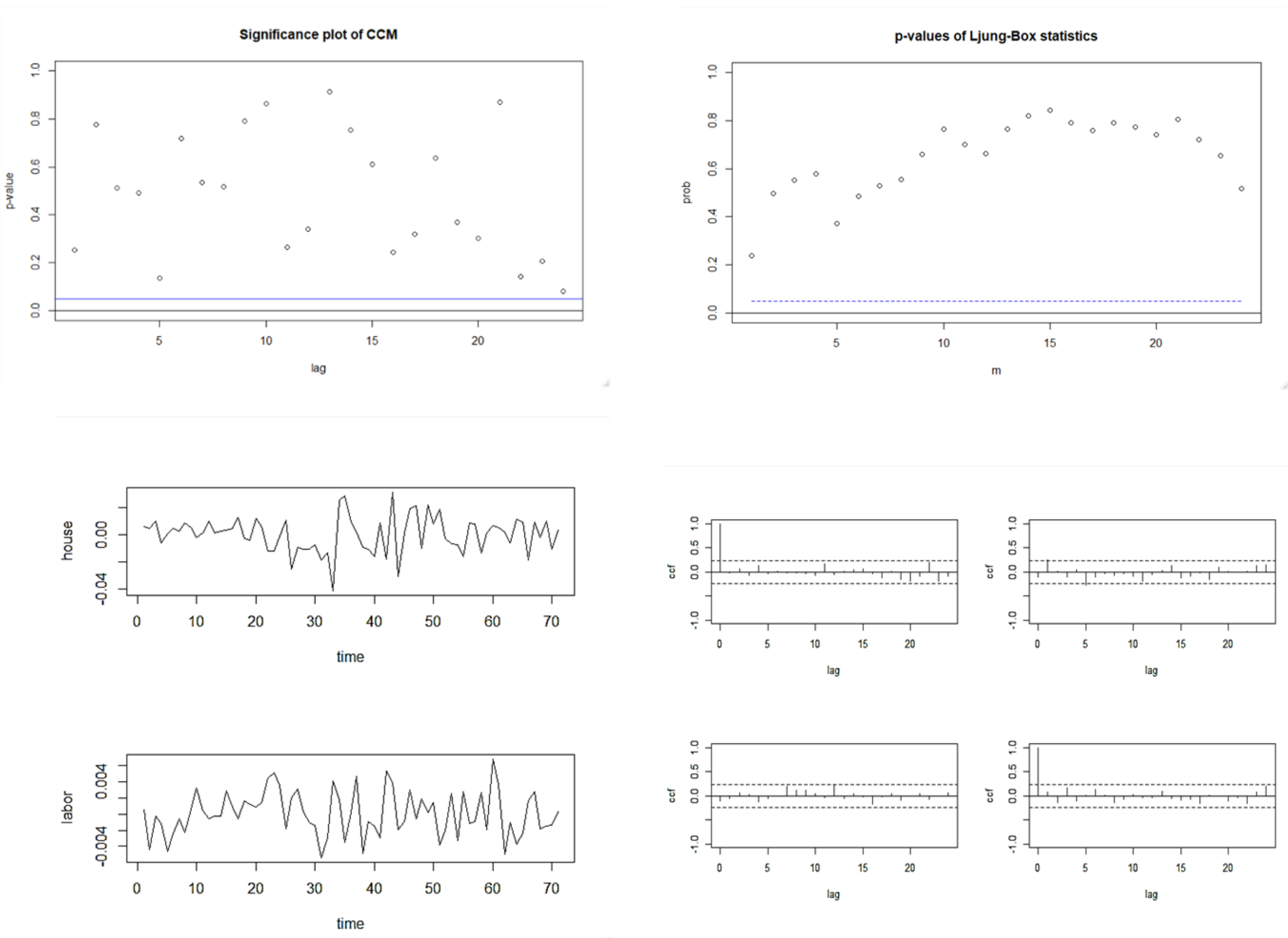


Figure 5: Model Checking for Model with Chicago Labor Force. From Top Left Clockwise – CCM Significance plot, Ljung-Box Statistics plot, Residuals Plot, CCF Plot

We also found the Impulse Response Function of this time series model, as shown in Figure 6. Observe from the (1,2) and (2,1) plot that their impulse responses are 0 for all time lags, agreeing with the results of the granger causality that there is no relationship between differenced Chicago housing prices and differenced Chicago labor force. The response of both

Chicago's housing prices and Chicago labor force to their own shock are large in magnitude, before steadily decreasing to 0 after several time lags.

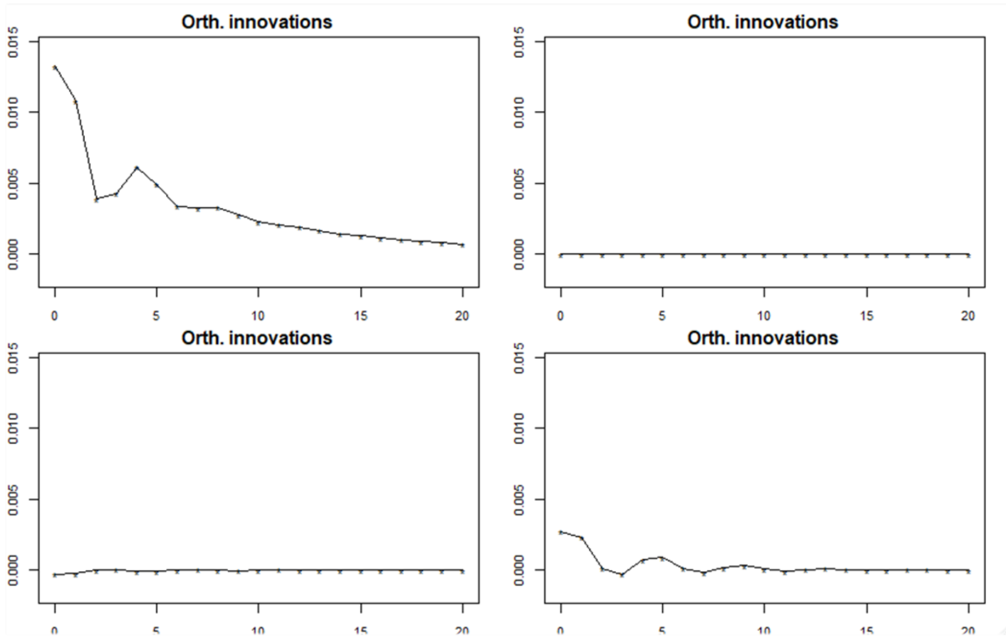


Figure 6: Impulse Response Function of Model 2 (Orthogonal, and non-cumulative).

From top left clockwise: Chicago to shock in Chicago, Chicago to shock in Labor, Labor to shock in Labor, Labor to shock in Chicago

Finally, we will use our model to perform forecasting and predict the movement of Chicago's housing prices (measured in index) after 2018 Q3, for up to 3 years, and simultaneously the forecast for Chicago's labor force as well.

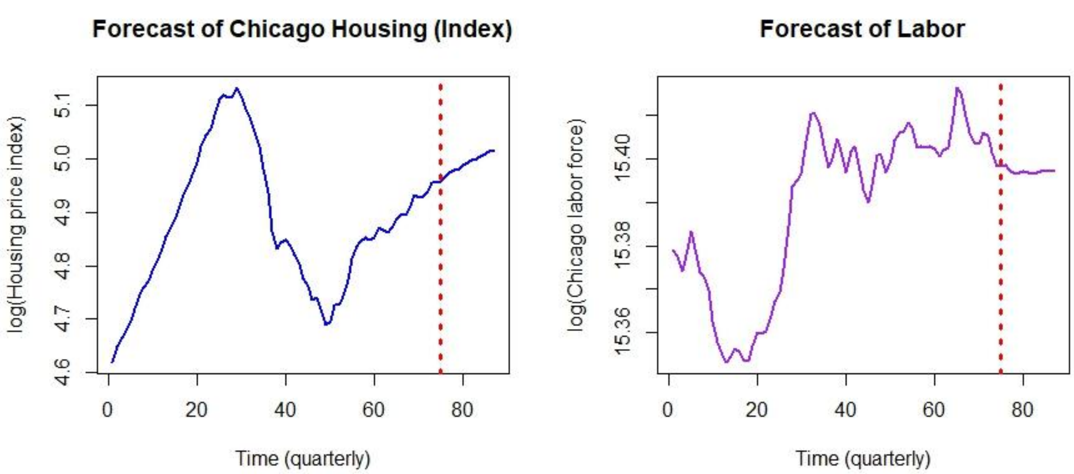


Figure 7: Forecast of Chicago Housing Price Index, Chicago Labor Force (Model 2)

Although there do not seem to be a relationship between Chicago housing prices and labor force, we proceed with the forecast in order to have a point of comparison with other models, and to illustrate the difference in forecasting when different information is available.

Since the model is expressed in ΔZ_t , we manually transform the predictions back to Z_t to have an easier comparison with other models. Note that the forecast is the time series after the red vertical line, as presented in *Figure 7*.

Observe from the plot the Chicago housing price index in this model is forecasted to continue to increase at a rather consistent rate, unlike that of model 1 (which has a sharp decrease after 2018 Q3 before increasing again). On the other hand, while the prediction does not directly capture the same concern of population decline in the literature review, it does not paint a rosy picture either – the Chicago labor force is forecasted to have a slight decrease before becoming relatively stagnant for the next 12 quarters.

4.3. Model 3 – Comparison with other Cities

Our last model involves a multivariate time series data consisting of 3 components, all of different cities' housing price index – Chicago (*CHI*), San Francisco (*SF*), and Detroit (*DT*), labelled as Z_1, Z_2, Z_3 respectively. The rationale for these choices is described in Section 2, and we aim to analyze to what extent an understanding of one city's change in housing prices affect another city. To begin, we plot both the logged time series data and first-order differenced time series, shown in *Figure 8*.

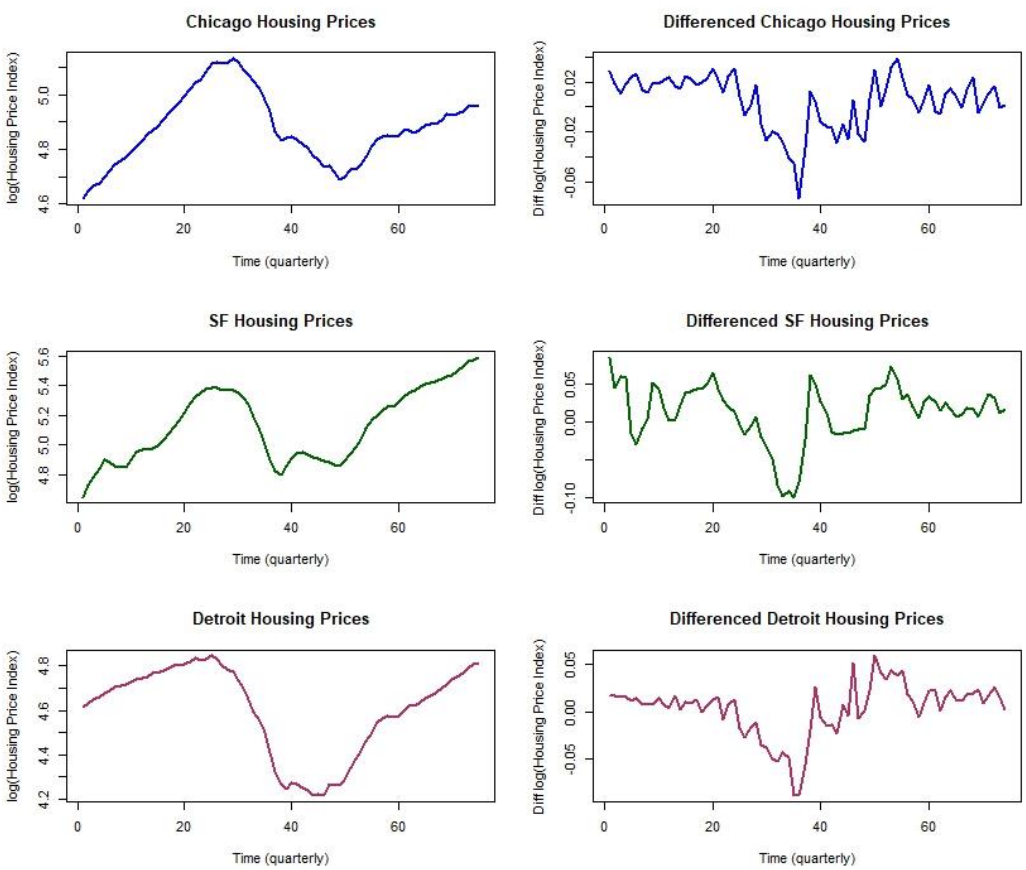


Figure 8: Time Series Plot of different cities, logged and differenced logged

Notice that in general, Detroit's price index takes on smaller values compared to Chicago, while San Francisco's index has the largest values. Before the recession, Detroit's housing prices grew at a smaller pace than both SF and Chicago, but post-recession, Chicago is now the city with the lowest average growth rate.

As before, we start off by checking the possibility of unit root. With the optimal AR order for each univariate time series, I performed the Augmented Dickey-Fuller test, obtaining a p-value of 0.78 0.760, and 0.622 for Chicago, SF and Detroit respectively, implying that unit root exists for all components.

We then found the optimal VAR order of 3 by the Chi-square test statistics, and performed the Johansen test for co-integration (including a constant term). However, we failed to reject the null hypothesis that r = 0, obtaining a test statistic of 20.24 < critical value of 22.00 at 5% significance level. This implies there is no co-integration between the 3 components. In order to eliminate the problem of unit root, we differenced the multivariate time series, with the new series depicted in *Figure 8*.

Subsequently, we check the differenced time series to ensure it does not contain any unit roots, and note that the Augmented Dickey-Fuller test returns p-values of 0.041, 0, and 0.012 respectively, implying that the issue of unit root has been resolved.

Before fitting a model, we check for Granger Causality between every pair of the 3 components using the Granger test function. The results are presented below:

Hypothesis	P-Value	Conclusion
Chicago granger-caused SF	0.0012	Yes
SF granger-caused Chicago	0.016	Yes
Chicago granger-caused Detroit	0.411	**No**
Detroit granger-caused Chicago	0.0005	Yes
SF granger-caused Detroit	0.0040	Yes
Detroit granger-caused SF	0.0001	Yes

We are able to establish granger causality in almost all the scenarios except the case where Chicago granger-caused Detroit – it implies that Chicago price index does not provide statistically significant information about future values of Detroit housing prices. This is surprising, given that Chicago and Detroit share many similarities in terms of locality and history. Perhaps, housing prices change in SF and Detroit are driven by some factors that are not present in Chicago, such as the former's active urban renewal and revitalization plans.

Nonetheless, we notice that Detroit and SF both granger-caused Chicago's housing price index. Hopefully, this will imply better forecasting for Chicago's housing prices.

We continue with fitting a VAR(3) model to the data, which gives us a model with an AIC of –26.09 and BIC of –25.25. In order to simplify the model and remove parameters that are not statistically significant, we refined the model with a threshold of 1.64, and obtained a new model with an AIC of –26.11 and BIC of –25.58. We proceed with the refined model that gives us a smaller BIC/AIC, described below.

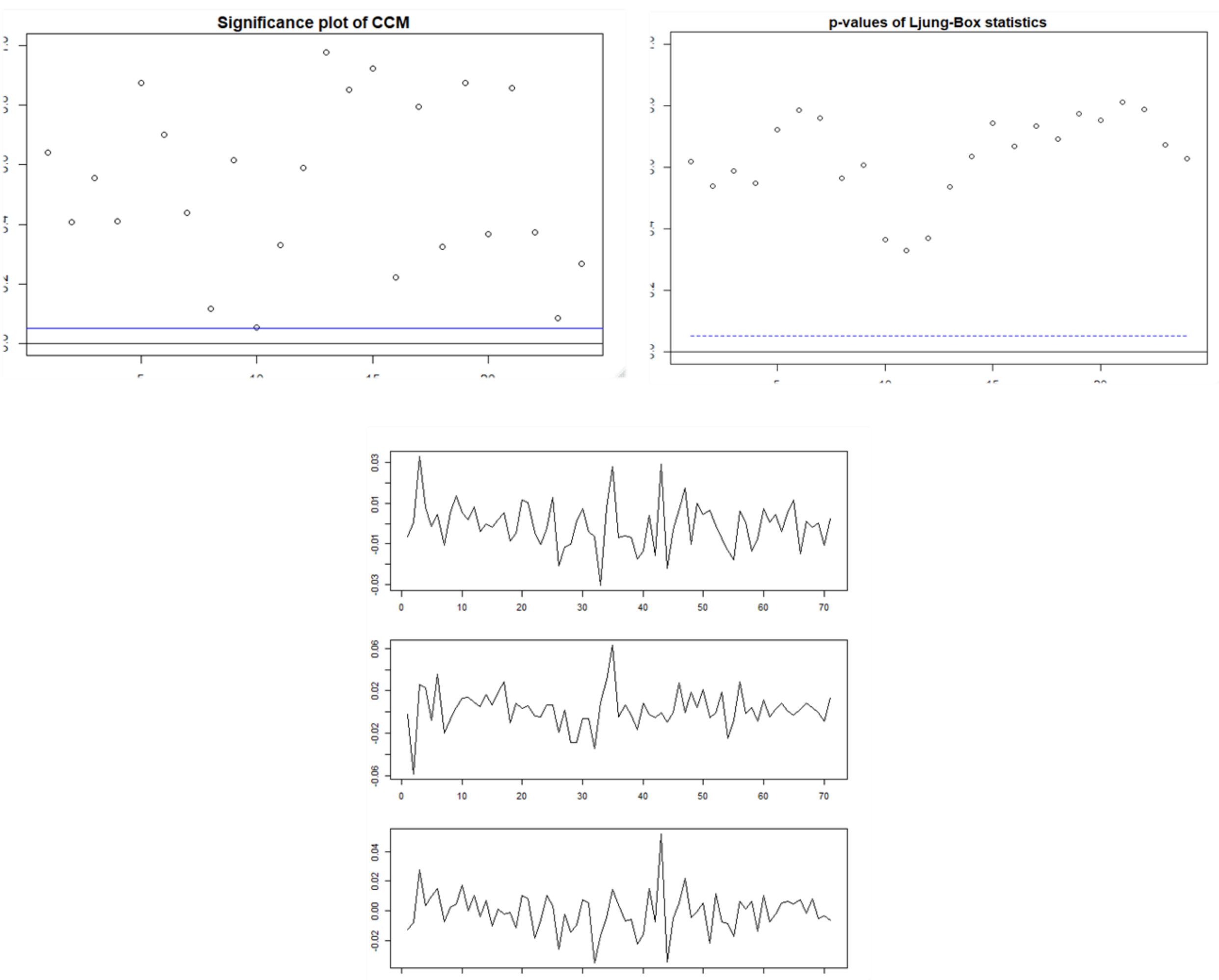
$$\Delta \boldsymbol{z}_t = \boldsymbol{\phi}_0 + \boldsymbol{\phi}_1 \Delta \boldsymbol{z}_{t-1} + \boldsymbol{\phi}_2 \Delta \boldsymbol{z}_{t-2} + \boldsymbol{\phi}_3 \Delta \boldsymbol{z}_{t-3} + \boldsymbol{a}_t$$

$$\boldsymbol{\phi_0} = \begin{pmatrix} 0 & 0 & 0 \end{pmatrix} \qquad \boldsymbol{\phi_1} = \begin{pmatrix} 0.366 & 0.396 & 0 \\ -0.425 & 1.280 & 0.251 \\ 0 & 0.523 & 0.365 \end{pmatrix}$$

$$\boldsymbol{\phi_2} = \begin{pmatrix} 0 & -0.202 & -0.002 \\ 0 & -0.567 & 0 \\ 0 & -0.394 & 0 \end{pmatrix} \qquad \boldsymbol{\phi_3} = \begin{pmatrix} 0.219 & 0 & 0.183 \\ 0.372 & 0 & 0 \\ -0.576 & 0.208 & 0.430 \end{pmatrix}$$

$$\boldsymbol{\Sigma_a} = 10^{-4} \begin{pmatrix} 1.286 & 0.811 & 1.091 \\ 0.811 & 3.035 & 0.967 \\ 1.091 & 0.967 & 1.843 \end{pmatrix}$$

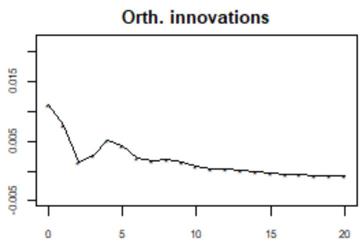
Residual Analysis suggests that the refined model is adequate.



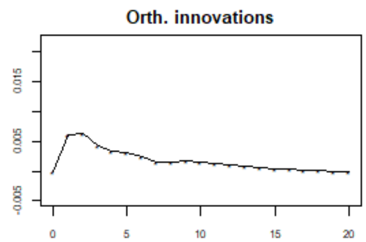
Residual Plots

Figure 9: Model Checking for Model 3

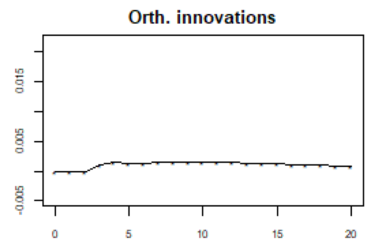
Once again, we calculated the Impulse Response Function (non-cumulative) of this time series model. Observe that for all components, the response of a component to a positive shock in another component is positive. Specifically, Chicago's response to a shock from SF has a delayed effect of 1 period, while the response to a shock from Detroit is also delayed and much smaller than other impulse responses.



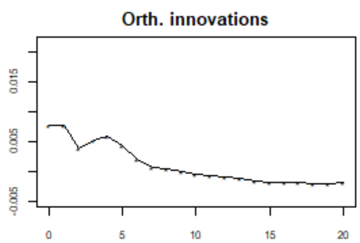
CHI to shock in CHI



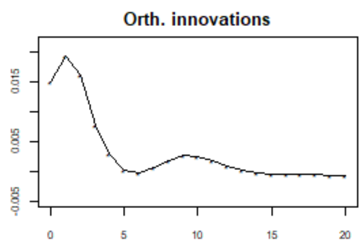
CHI to shock in SF



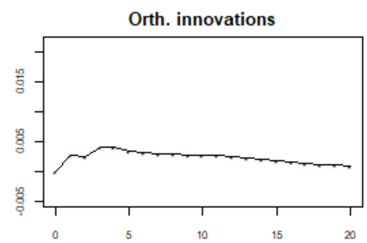
CHI to shock in DT



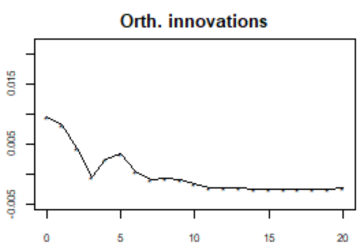
SF to shock in CHI



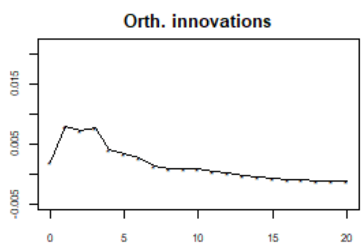
SF to shock in SF



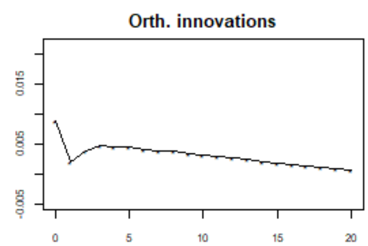
SF to shock in DT



DT to shock in CHI



DT to shock in SF



DT to shock in DT

We conclude our third model with forecasting to predict the movement of Chicago's housing prices (measured in index) after 2018 Q3, for up to 3 years, and simultaneously the forecast for both San Francisco and Detroit. The predictions of ΔZ_t are listed in *Table 1*. Like in model 2, we manually transform the predictions from ΔZ_t back to Z_t to allow for easier comparisons.

Table 1: Forecast of Marginal Change in Chicago, SF, Detroit Housing Price Index (ΔZ_t)

Horizon	1	2	3	4	5	6	7	8	9	10	11	12
CHI	.0109	.0113	.0045	.0047	.0063	.0059	.0040	.0035	.0038	.0035	.0026	.0021
SF	.0215	.0161	.0089	.0069	.0079	.0069	.0052	.0046	.0043	.0035	.0024	.0017
DT	.0135	.0182	.0100	.0059	.0069	.0074	.0051	.0028	.0027	.0024	.0011	.0001

Observe that the forecast for all 3 cities follow the same trend of their housing price index increasing, albeit at a slower rate than before the forecast origin. Detroit's forecast predicts that their housing price will increase, eventually above the previous high point attained before the 2008 housing crisis, while San Francisco has already gone past its previous high point a few years before 2018. However, we did not observe the same in Chicago's forecast, as growth is predicted to slow, without going past the previous record of ~5.15 (log index) at the end of the 3 years forecast. This agrees with market sentiments of a slowdown, and the belief that Chicago's housing markets are not expected to hit peak values until 2020 (Murray & Schuetz, 2018). Indeed, note from *Table 1* that Chicago has the smallest predicted values for ΔZ_t for the first few horizons, which measures the marginal change to the city housing price index for each city, reinforcing the idea that in the aftermath of the Housing Crisis, Chicago has a slower recovery compared to San Francisco and Detroit.

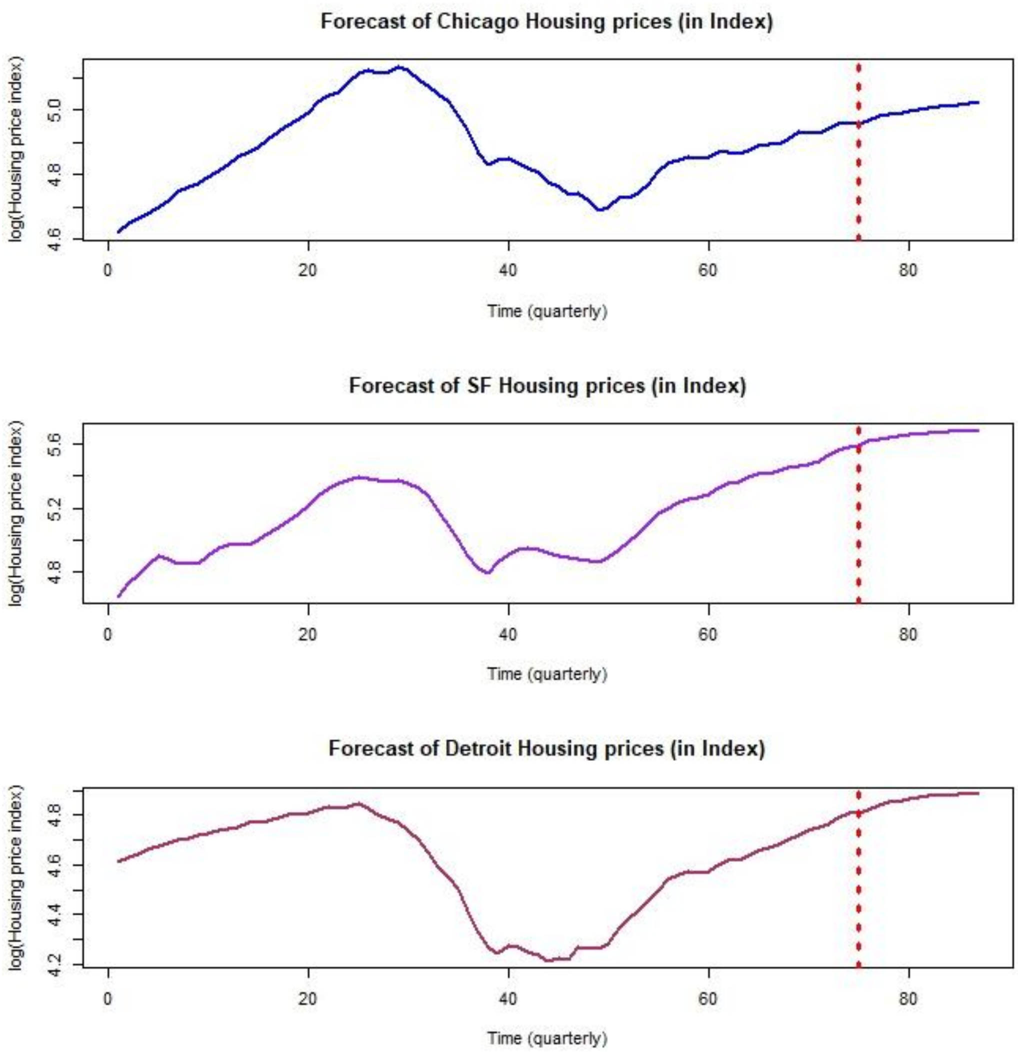


Figure 10: Forecast of Chicago, SF, Detroit Housing Price Index (Model 3)

4.4. Forecast Comparison

Let us compare the 3-year predictions for the different models, with 2018 Q3 as the origin. The values below are logged price index. Observe that we have 3 forecasts of Chicago's housing prices from the 3 different models.

Table 2: Summary of Forecast Prediction from 3 models

Horizon		1	2	3	4	5	6	7	8	9	10	11	12
Model 1	CHI	4.964	4.908	4.841	4.813	4.804	4.804	4.807	4.814	4.822	4.830	4.838	4.846
Model 2	CHI	4.967	4.975	4.979	4.983	4.988	4.994	4.998	5.001	5.006	5.010	5.014	5.017
Model 3	CHI	4.969	4.980	4.985	4.989	4.995	5.001	5.005	5.008	5.012	5.016	5.018	5.020
Model 3	SF	5.613	5.629	5.638	5.644	5.652	5.652	5.659	5.669	5.673	5.677	5.680	5.681
Model 3	DT	4.822	4.823	4.851	4.857	4.864	4.871	4.876	4.879	4.882	4.884	4.885	4.886

Predictions from model 2 and model 3 are close to each other in terms of the numerical values and rate of change. However, results from model 1 are unlike model 2 and 3's results. One possible reason is because of the fundamental differences in the multivariate time series data and model employed. In the first model, we performed error correction since there exist co-integration, and hence the prediction is made in the form of Z_t. In model 2 and 3, since we failed to find any co-integration, we proceeded with a model utilizing the first order differenced time series, before transforming the prediction to the Z_t scale. This fundamental difference may explain why the forecasts vary.

Next, in order to compare between model 2 and 3, we turn to the standard deviation of the forecast error, referring specifically to the component of Chicago housing price index.

Table 3: Standard Deviation of Forecast Errors from Model 2 and 3 for $\Delta Z_{t,1}$ (Chicago's Price Index)

Horizon	1	2	3	4	5	6	7	8	9	10	11	12
Model 2	.0133	.0171	.0175	.0181	.0191	.0197	.0200	.0202	.0205	.0207	.0208	.0209
Model 3	.0113	.0147	.0165	.0175	.0185	.0192	.0195	.0198	.0201	.0203	.0204	.0206

Note from the plot that for all horizons, the standard deviation of the forecast errors is smaller for model 3 than for model 2. Recall that in model 2, no granger causality is present between the 2 components, and the IRF reflects this as well. On the other hand, in model 3, granger causality exists for the majority of the component pairs. Hence, the additional information available in the past values of San Francisco's and Detroit's housing prices helps to reduce the uncertainty in the prediction of Chicago's future housing prices.

5. Conclusion

In this paper, we investigated macroeconomic factors affecting Chicago's housing prices, and compared the performance of housing markets across different US cities, with the ultimate goal of improving the forecast for Chicago's housing prices measured via the S&P Price Index.

The results from model 1 and model 2 describe how national housing supply and Chicago labor force affects Chicago's housing prices, measured in the Housing Price index.

In the first model, we analyzed the bivariate time series data consisting of National Housing Price Index, Chicago Housing Price Index, and National Housing Supply. Due to co-integration, we fitted an VAR error correction model, and observed that national supply has a significant impact on the Chicago housing price index, both in terms of granger causality and by the impulse response function values. In the IRF, given a positive shock in national housing supply, there is an initial positive response in Chicago's housing price index to that shock, before it eventually converges towards a negative value, indicating that in the long-term, if housing supply increases, Chicago's housing price is likely to decrease. This supports the findings by Glaeser and Gyourko (2008) that housing permits and housing prices are negatively correlated.

In our second model, we analyzed a bivariate time series data consisting of Chicago's housing prices and Chicago's labor force. Surprisingly, there turns out to be no granger causality in either direction for the differenced time series data, and the lack of relationship is also supported by the parameters in our refined model and the impulse response functions. This findings for Chicago do not agree with Johnes and Hyclak's (1996) discoveries in Houston, Milwaukee, Fort Lauderdale, and Hartford, where they observed that there was a bidirectional relationship between the labor force and city housing prices in all 4 cities. One explanation is that the Chicago Civilian Labor Force measures the entire metropolitan area's (Chicago–Joliet–Naperville) labor force, while the housing price index measures only properties within the City of Chicago. A possible extension is to utilize either a more specific geographic labor force indicator, or to turn to metrics measuring the employment rate.

In our third and final model, we analyzed the multivariate data consisting of housing price index from Chicago, San Francisco, and Detroit, and fitted a VAR model to the differenced data since there exists no co-integration. Granger Causality exists between most of the paired components, and most of the Impulse Response Functions are positive as well. The forecast results are similar to model 2's forecast, except with a lower forecast error standard deviation for Chicago's housing price index, suggesting that Detroit's and San Francisco's price index provide information for better forecasting. From the forecasted trends, we also observed that Chicago is projected to have slower housing market growth compared to San Francisco and Detroit, agreeing with the literature.

- END -

References

1. Tsay, R. S. (2013). *Multivariate Time Series Analysis: With R and Financial Applications*. John Wiley & Sons.
2. Glaeser E.L, Gyourko, J. (2008). *Rethinking Federal Housing Policy*. The AEI Press, Washington D.C.
3. Mankiw, N.G., Weil, D.N. (1989, Jan). *The Baby Room, The Baby Bust, and the Housing Market.* Regional Science and Urban Economics, 19, 235-258.
4. Johnes, G., Hyclak, T. (1999). *House Prices and Regional Labor Markets*. The Annals of Regional Science, 33(1), 33-49.
5. Capozza, D.R., Hendershott, P.H., Charlotte, M., Mayer, C.J. (2002). *Determinants of Real House Price Dynamics*. National Bureau of Economic Research Working Paper 9262.
6. Davidoff, T. (2006, Mar). Labor income, housing prices, and homeownership. *Journal of Urban Economics*, 59(2), 209-235.
7. Murray, C., Schuetz, J. (2018, Jun). *Housing in the US.* The Brookings Institution Research.
8. Hill, B. (2019, Apr 17). *All Illinois Metro Areas saw Population Shrink in 2018*. Illinois Policy.
9. Divounguy, O., Hill, B. (2018, Dec 06). *Chicago Anticipated to be Nation's Weakest Major Housing Market in 2019*. Illinois Policy.
10. Dvorkin, M.A., Shell, H.G. (2016). *The Recent Evolution of U.S. Local Labor Markets.* Federal Reserve Bank of St. Louis Economic Research, Economic Synopses, 2016 No. 15.
11. Federal Reserve Bank of St. Louis and US. (2019). *FRED Economic Data*. Office of Management and Budget.