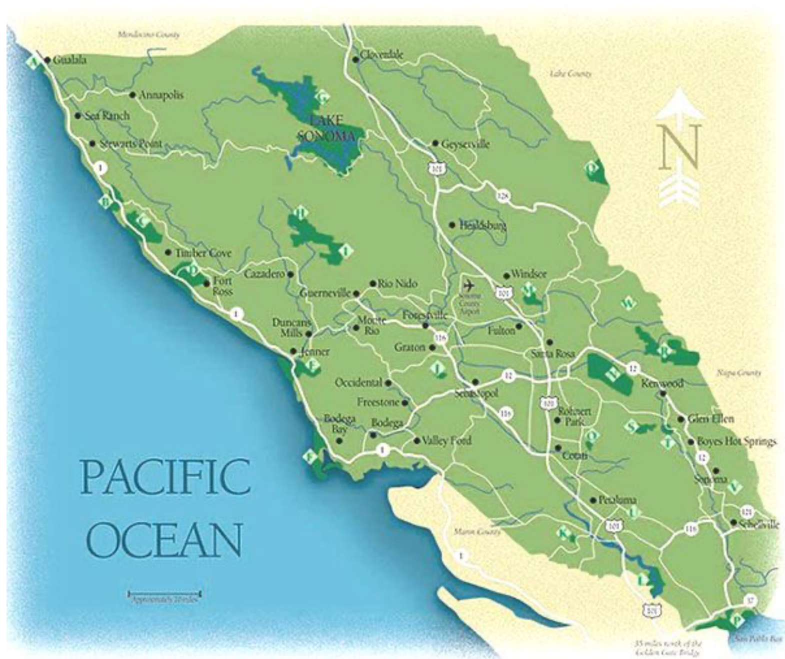
A Financial Time Series Analysis of Solano and Sonoma Counties



Ella Archie, Jin Burge, Shayan Golshan, Alex Mitroff, Jadin Stewart

GSB S570 - 01

June 15, 2023

Table of Contents

A. Overview of Two Counties' Housing Market..2
B. One Application for the Econometric Analysis.. 4
C. Properties of the House Price Time-Series..5
D. ARIMA models of prices and returns...8
E. Forecasting Power of the ARIMA model(s)... 9
F. Multivariate analysis..10
G. Conditional Variance Analysis: Various types of GARCH models................................... 11
H. Value-At-Risk Analysis...14
I. Conclusion and Managerial Implications...15
Works Cited.. 16
R Code.. 17

A. Overview of Two Counties' Housing Market

Sonoma County

Sonoma County, known for its coastal landscape, wine country, and tourism, is located about an hour northwest of the San Francisco Bay Area. Sonoma County accounts for 1,575 square miles of California and is home to 487,001 individuals ("Sonoma County Tourism Office Site"). The majority of residents in this county fall between the ages of 25 and 64 with the median age being 42.1. It is also one of the wealthiest counties in California with 59% of the population earning above 300% of the Federal Poverty Line making housing more expensive than many other places ("Demographics and Social Characteristics").

Over the past 3 years, the housing market in Sonoma County has experienced a steady upward trend. In April 2020, realtor.com reported the median listed home price to be $699,000 and the median home price sold to be $627,000. In March of 2023, the median listed home price had increased to $829,000 and the medium home price sold to $789,900. Over the span of three years, the median listed home price increased by roughly 18% and the median home went from selling at roughly 10% below the list price to only 4% below the list price. As shown by these numbers, Sonoma County's housing market is becoming more and more competitive with demand for housing exceeding supply.

While the housing market in the greater San Francisco Bay Area has seen a downward trend in recent years, Sonoma County has become an outlier as it continues to see a rise in the market. The North Bay Business Journal suggests two reasons why Sonoma County is defying the downward trend: inventory and desirability. First, the low inventory (number of homes listed for sale) in the County coupled with the large number of individuals looking for housing is creating a vast demand and inflating prices. Second, with companies allowing more and more employees to work from home, individuals have demonstrated they are willing to live far away from their company headquarters if it means living in a nicer area. Sonoma County is beautiful and so it is no surprise that Bay Area work-from-home professionals have found themselves moving out West.

Solano County

Located almost directly between San Francisco and Sacramento, Solano County spans seven cities, most notably Fairfield, Vacaville, and Vallejo. The county spans 909.4 square miles, including 84.2 square miles of water area, and 675.4 square miles of rural land. Solano is the 19th most populated county in California out of 58 counties, with a population of about 451,432 people as of the 2020 census counts.

In terms of the housing market, Solano County has had a fairly consistent upward trend in the housing market, with a small downturn starting in June 2022. As of March 2023, Zillow average home values are $570,922, down 2.8% from 2022. Like the vast majority of California, the housing market is being impacted by a shortage of housing, especially affordable housing. The low supply of housing is exacerbated by high demand from buyers, resulting in higher home prices and competition within the buyer's market. This has resulted in bidding wars, with sellers generally fielding multiple offers above the listed asking price. In response, cities such as Vallejo and Fairfield have increased the amount of new construction in an attempt to accommodate for the high volume of demand.

The slight downturn that the Solano County housing market has experienced since June of 2022 reflects a larger trend in the state of California. While housing prices are still significantly inflated, the market as a whole has taken a slight dip. Rising interest rates, and the expectation that they will only continue to increase, coupled with some economic uncertainty, especially in the tech sector, likely contributed to this decline. Sellers are also likely to be discouraged from selling because in doing so they forfeit their current mortgage rates which are likely to be significantly lower than the new rates, incentivizing them to hold on to the property (and their existing mortgage rates) rather than selling and purchasing a new home. Despite these influences, the market remains strong and continues attracting buyers and investors worldwide.

B. One Application for the Econometric Analysis

Clients who would be interested in seeing the differences between Solano County and Sonoma County would include but are not limited to, homeowners and those interested in moving to the area, local government officials, as well as real estate investors, agents, and brokers. All of these groups could get benefits from understanding the econometric analysis of the housing markets of Sonoma and Solano counties.

Homeowners are also able to benefit if they are interested in understanding the current state of the housing market. Having this information readily at hand could help them drive decisions relevant to buying or selling (or holding) property in either county. Next, government officials would most likely benefit from this econometric analysis and find interest in the impact of new regulations/policies that might arise within each county. These impacts may help guide land use and sites for future home development. Real estate agents/brokers, may utilize our econometric analysis to develop strategies to figure out the correct market segment. Lastly, real estate investors would benefit greatly from these insights since they would have an idea of where their money is going. Being able to create an informed decision on where to invest as well as the trends associated with prices is invaluable in the long run. With this much capital at stake, real estate investors must do their due diligence or they risk losing assets.

C. Properties of the House Price Time-Series

Table 1: Descriptive Statistics

	Mean	Std Dev	Median	Min	Max	Kurtosis
Sonoma	158.06	90.69	135.07	38.2	394.46	-0.56
Solano	158.06	102.09	154.41	32.44	414.32	-0.94

Figure 1 illustrates a line plot of the housing prices in Solano and Sonoma counties overtime. The graph reveals that the prices between the two counties have remained fairly similar over time, following comparable trends and patterns throughout the years. Housing prices have steadily increased since 1980, with a local maximum around 2008.

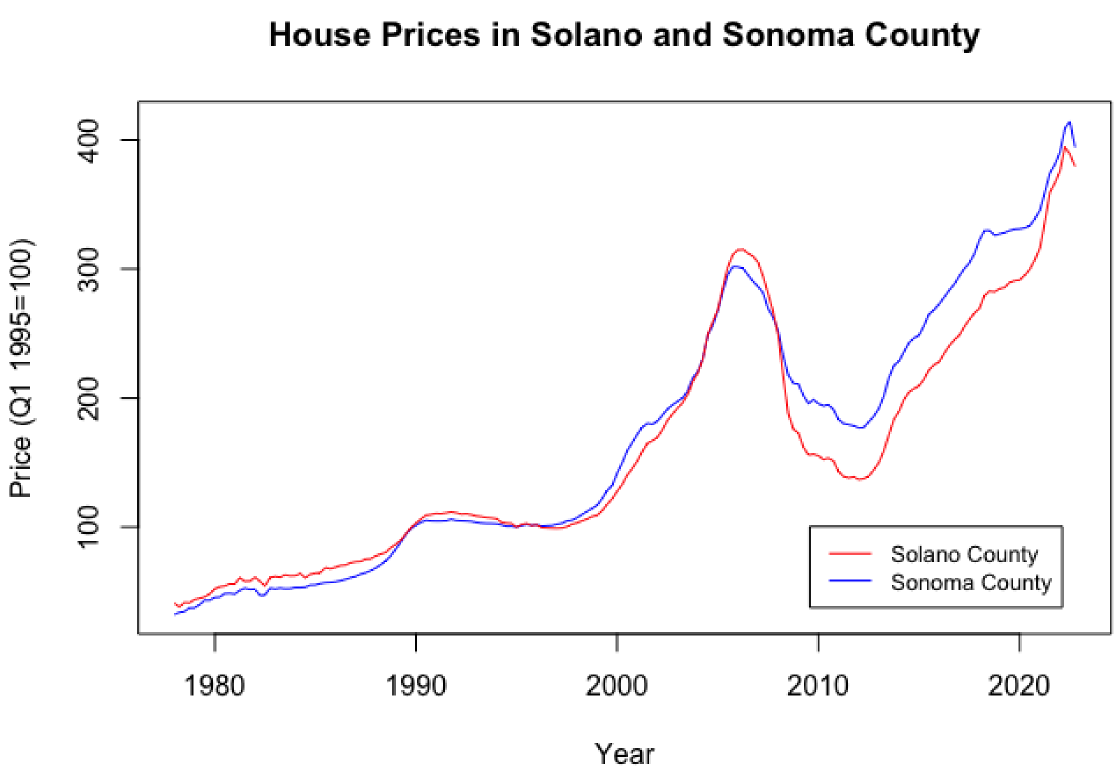


Figure 1: House Prices in Solano and Sonoma County

The Augmented Dickey-Fuller Test on the original prices resulted in considerable p-values, indicating that the original prices are unit root and nonstationary. The Augmented Dickey-Fuller Test on the log prices resulted in small p-values, indicating that the log prices are stationary.

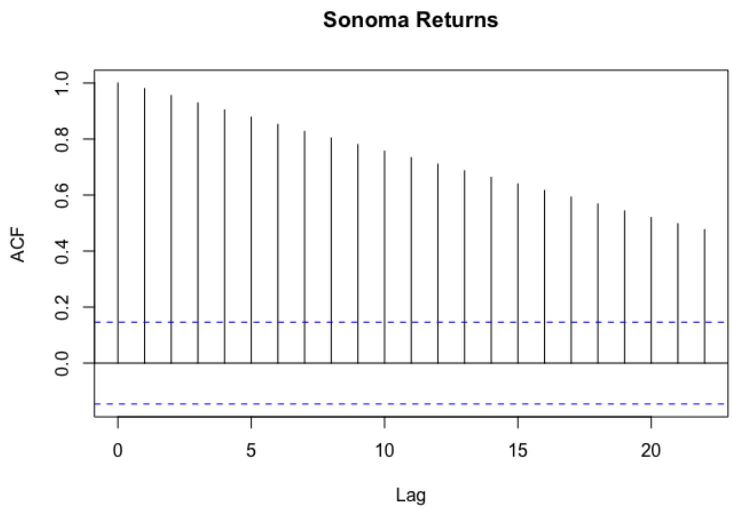


Figure 2: ACF Graph of Sonoma Returns

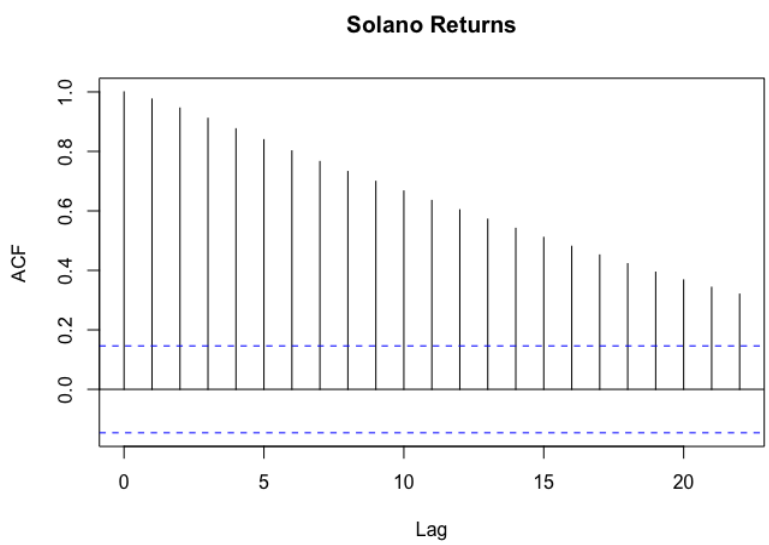


Figure 3: ACF Graph of Solano Returns

Table 2: Augmented Dickey-Fuller Test

	Test Stat	**P-value**
Sonoma - Orig Prices	-3.0423	0.1406
Solano - Orig Prices	-2.9461	0.1808
Sonoma - Log Prices	-3.7968	0.02067*
Solano - Log Prices	-3.526	0.0419*

Additionally, the ACF graphs were created for both counties, as illustrated in Figures 2 and 3. Both Sonoma and Solano County have positive autocorrelations, revealing that at a given time period, it is likely to be correlated with the values at a previous time period. Since it gradually decreases, the correlation gets weaker as the lag increases. All lags have some statistical significance and the decreasing ACF gives hints to make use of the ARIMA model. The Autoregressive Component captures the relationship between the current value in the time series and the lagged values and the Moving Average Component captures the relationship between a current value and the residual errors. The Integration Component accounts for non-stationarity in the data by applying differencing.

Table 3: Seasonality Test Using SeaDum Test

	Test Stat	**P-value**
Sonoma - Orig Prices	1.09	0.3554846
Solano - Orig Prices	0.5	0.6818125
Sonoma - Log Prices	1.18	0.3198172
Solano - Log Prices	0.3	0.8261503

Next, a seasonality test was conducted using the SeaDum test, which utilizes dummy variables. The results revealed large p-values, all above statistical significance, illustrating that there is no seasonality with either country regardless if the prices are original or log prices.

D. ARIMA models of prices and returns

ARIMA models allow us to better understand how the level of returns is dependent on past (lagged) values, and thereby enable us to create forecasts of future returns. We have previously demonstrated that the housing price returns are stationary (obtained by getting the log first difference of the index) and that none of the time series exhibit seasonality. Therefore we will fit an ARIMA model (using the auto.arima function from the *forecast* R package) without enabling seasonality. The resulting models fit are as follows:

Table 4: ARIMA Model Results

County	**ARIMA Autofit Model**	**Accuracy (RMSE)**	**AIC**
Sonoma Housing Prices	ARIMA(3,1,1)	4.356	1029.54
Solano Housing Prices	ARIMA(5,1,2)	5.529	1120.2
Sonoma Log Returns	ARIMA(5,0,3) with zero mean	0.0219	-827.98
Solano Log Returns	ARIMA(3,0,4) with zero mean	0.0278	-747.46

Additionally, we explore a few ARIMA-X models by introducing exogenous variables which may have predictive power on housing prices: interest rates, unemployment rates, housing permits, and real GDP are brought in using our FRED API. All of these variables pass our tests for both stationarity and seasonality, except interest rates and GDP which we use as a first-differenced time series. As we might expect, in both counties the coefficients on interest rates and unemployment rates are both negative, reflecting a negative correlation between housing prices and past levels of these variables. Housing permits and GDP do not induce significant effects when all four variables are included, but we must concede that these data sets are perhaps too broad (being a measurement of the entire United States rather than a more specific region or MSA) to have adequate predictive power in our model. The AIC values for these models are slightly better than we saw with no exogenous variables (e.g. Sonoma returns: -831 vs. -827), and we can further improve this fit by excluding the two less relevant variables mentioned previously. Moving forward, we will be limited in our ability to use these ARIMA-X models for forecasting, due to that requiring us to reliably predict the external regressors which is beyond the scope of this project.

E. Forecasting Power of the ARIMA model(s)

Having fit the ARIMA models above, we can produce forecasts of both housing prices and returns. These results are displayed on the plots below (see Figure 4). We observe stagnant or even declining returns in both counties. The nominal prices predict stagnant growth in Sonoma County, but a decline (or mean reversion) in Solano County.

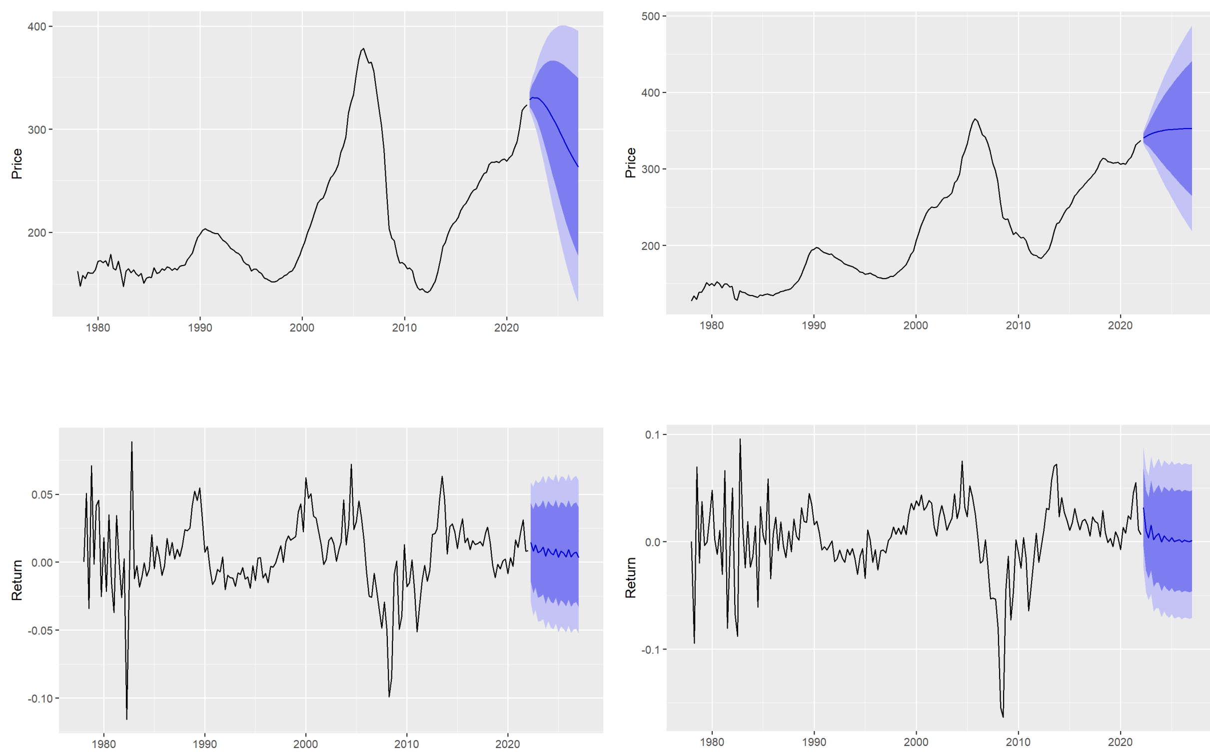


Figure 4: Forecasts of Sonoma and Solano Prices and Returns

F. Multivariate analysis

In regards to the multivariate analysis, a VAR model was created rather than VECM because the cointegration test observed that one county was zero rank and other was full rank. VECM should only be applied if a middle value is present. This reveals that a linear combination of two series is already stationary since it is full rank. The VAR model reveals that past values of Sonoma affect the current prices of Solano, and vice versa. The same conclusion was found for the returns, revealing that past returns of Sonoma affect the current returns for Solano, and vice versa.

Table 5: Cointegration of Sonoma and Solano Prices

	test	**10pct**	**5pct**	**1pct**
r <= 1	0.33	6.50	8.18	11.65
r = 0	9.38	12.91	14.90	19.19

Table 6: Cointegration of Sonoma and Solano Log-Prices

	test	**10pct**	**5pct**	**1pct**
r <= 1	21.10	6.50	8.18	11.65
r = 0	82.29	12.91	14.90	19.19

We also conducted a Granger Causality test, illustrating that Solano prices do Granger-cause Sonoma prices, and Sonoma prices do Granger-cause Sonoma prices.

For returns, the results indicate that Sonoma returns do not Granger-cause Solano returns, but Solano returns Granger-cause Sonoma returns.

Table 7: Granger Causality

	F-Stat	**P-value**
Sonoma - Nominal	4.0994	0.002951**
Solano - Nominal	3.8646	0.004394**
Sonoma - Returns	1.0466	0.4027
Solano - Returns	4.0301	6.881e-05***

G. Conditional Variance Analysis: Various types of GARCH models

The Generalized Autoregressive Conditional Heteroscedasticity (GARCH) model is proving popular today for analyses of the housing market. The goal in fitting a GARCH model is to see if we can anticipate the volatility of housing returns in our two markets of interest. Specifically, these models can show us the *significance* of the variance of returns (whether it exceeds the unconditional variance of the distribution) and how it evolves over time. Additionally, GARCH can fit an AR process to the returns, which will come into play as we consider different models.

The first model considered is a GARCH(1,1), which will provide a useful 'baseline' for this analysis. Fitting GARCH(1,1) to the real returns of both counties, we are able to immediately see a few insights about the historical level of volatility:

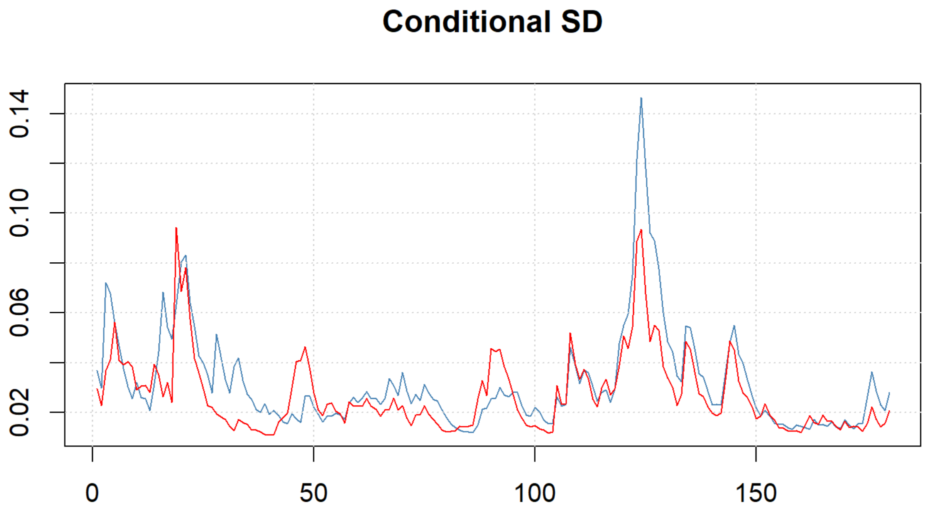


Figure 5: Conditional SD Plot of Volatility

Coefficients on mu, alpha1, and beta1 are significant at the 5% level or lower in both counties. Therefore we confirm the presence of GARCH effects. However, which variant of GARCH model proves to be the best fit for modeling volatility? We augment our analysis by considering two variants: IGARCH and TGARCH. The methodology used is as follows: examine the distribution of standardized residuals for normality (low skewness, low kurtosis); test both standardized residuals and squared residuals for autocorrelation with a Box-Ljung test; evaluate all models using Akaike Information Criterion and select lowest value. The results of these tests are given below, identifying the IGARCH(1,1) as the best fit model for further analysis.

Table 8: GARCH Model Results

	GARCH(1,1)		IGARCH(1,1)		TGARCH(1,1)	
	Sonoma	Solano	Sonoma	Solano	Sonoma	Solano
Skewness	-0.367	-0.311	-0.045	-0.117	0.094	-0.117
Kurtosis	0.389	0.305	1.022	0.939	0.726	0.685

AIC	-4.584	-4.269	**-5.125**	**-4.675**	-5.113	-4.605
Box test SR	REJECT	REJECT	**ACCEPT**	**ACCEPT**	ACCEPT	REJECT
Box test SSR	ACCEPT	ACCEPT	**ACCEPT**	**ACCEPT**	ACCEPT	ACCEPT

Once we have selected the IGARCH(1,1) for modeling volatility, it is interesting to note the historical level of volatility against returns. R produces some useful plots for this aim shown below in figures 6 and 7. We should exercise caution when forecasting volatility that we are not currently in the midst of a historically very high or very low period.

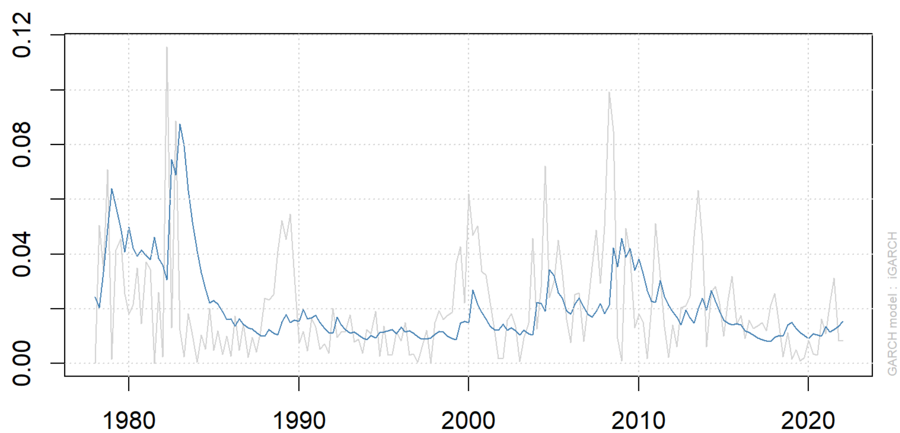


Figure 6: Sonoma Returns vs Volatility

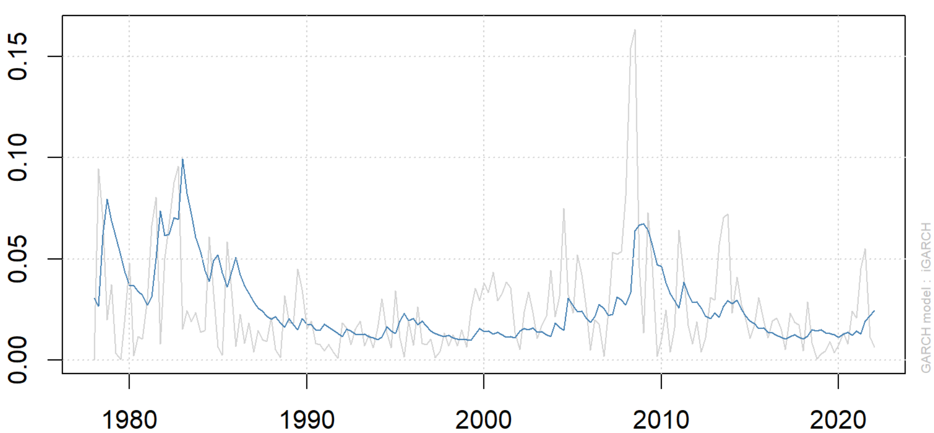


Figure 7: Vallejo Returns vs Volatility

H. Value-At-Risk Analysis

	Value at Risk (VaR)	Expected Shortfall (ES)
Risk Metrics Method	3.26%	4.09%
Econometric (GARCH) Modeling	2.04%	2.98%
Empirical Quantile	4.71%	N/A

Table 9: Value at Risk and Expected Shortfall - Sonoma

	Value at Risk (VaR)	Expected Shortfall (ES)
Risk Metrics Method	4.51%	5.66%

Econometric (GARCH) Modeling	3.45%	4.69%
Empirical Quantile	5.23%	N/A

Table 10: Value at Risk and Expected Shortfall - Solano

For Sonoma, at an $829,000 position (2023 median listed home price), the maximum VaR is 4.71% resulting in a maximum loss of $39,024.34. The maximum expected shortfall is 4.09% resulting in a maximum loss of $33,864.92. For Solano, At a $570, 922 position (the 2023 mean listed home price), the maximum VaR is 5.23% resulting in a maximum loss of $29,882.87. The maximum expected shortfall is 5.66% resulting in a maximum loss of $32,306.28.

I. Conclusion and Managerial Implications

From our ARIMA models, we expect stagnant if not slightly negative returns for the coming 10 periods in both counties. The upper 95% bound of our forecasted returns in the first quarter is 3.52% for Sonoma and 5.44% for Solano, just barely exceeding the 95% VaR for these two counties and falling below the 95% expected shortfall. Therefore, now is not the time to buy real estate expecting a return on investment.

Based on the forecasts, the prices appear to be decreasing. This will result in significant losses for houses that are not going to increase in profit. Returns are increasing slightly, however, overall the housing index is appearing more negative. Thus, this is a poor investment and much risk is involved. Selling now and buying later is the recommended course of action.

Works Cited

"Demographics and Social Characteristics." *County Of Sonoma*, https://sonomacounty.ca.gov/health-and-human-services/health-services/about-us/demographics-and-social-characteristics.

"Sonoma County Tourism Official Site." *Sonoma County Tourism Official Site | SonomaCounty.com*, https://www.sonomacounty.com/.

"Sonoma County, CA: Realtor.com®." *Realtor.com*, https://www.realtor.com/realestateandhomes-search/Sonoma-County_CA/overview.

U.S. Census Bureau. "QuickFacts: Solano County, California." Census Bureau QuickFacts, 2021, https://www.census.gov/quickfacts/solanocountycalifornia.

Wood, Susan. "Why Home Prices in These 4 Local Neighborhoods Are Bucking Bay Area Declines." *The North Bay Business Journal*, The North Bay Business Journal, 5 Feb. 2023, https://www.northbaybusinessjournal.com/article/article/why-home-prices-in-these-4-sonoma-county-areas-are-defying-bay-area-downwar/.

World Population Review. "Solano County, California Population 2021 (Demographics, Maps, Graphs)." World Population Review, 2021, https://worldpopulationreview.com/us-counties/ca/solano-county-population.

Zillow Group. "Solano County Home Prices & Values." Zillow, 2021, https://www.zillow.com/home-values/1395/solano-county-ca/.

R Code

(**all code and output can also be found in the html file attached in submission**)

````
```{r setup, include=FALSE}
knitr::opts_chunk$set(echo = TRUE)
#suppressMessages(library(gt))
#suppressMessages(library(gtExtras))
suppressMessages(library(remotes))
suppressMessages(library(fredr))
#suppressMessages(library(ggplot2))
#suppressMessages(library(tidyverse))
suppressMessages(library(tseries))
suppressMessages(library(stats))
suppressMessages(library(forecast))
#suppressMessages(library(stargazer))
suppressMessages(library(Quandl))
suppressMessages(library(quantmod))
suppressMessages(library(psych))
suppressMessages(library(e1071))
suppressMessages(library(MTS))
suppressMessages(library(xts))
suppressMessages(library(lmtest))
suppressMessages(library(seastests))
suppressMessages(library(urca))
suppressMessages(require(fGarch))
````
````

```
suppressMessages(require(rugarch))
suppressMessages(require(vars))


fredr_set_key('e6203f9b43cab4fa3302c61589709bc2')
start = "1978-01-01"
end = "2022-10-01"
vallejo_api = 'ATNHPIUS46700Q'
santarosa_api = 'ATNHPIUS42220Q'
intrate_api = 'FEDFUNDS'
unrate_api = 'UNRATE'
gdp_api = 'GDPC1'
permit_api = 'PERMIT'
cpi_api='CPALCY01USQ661N'
```
```

```
Obtaining Housing Data

```{r}
sonoma = fredr(
  series_id = santarosa_api,
  observation_start = as.Date(start),
  observation_end = as.Date(end)
)

vallejo = fredr(
  series_id = vallejo_api,
  observation_start = as.Date(start),
  observation_end = as.Date(end)
)

cpi = fredr(
  series_id = cpi_api,
  observation_start = as.Date(start),
  observation_end = as.Date(end)
)

sonoma_price = sonoma$value
vallejo_price = vallejo$value
sonoma_real = (sonoma_price/cpi$value)*100
```

```
vallejo_real = (vallejo_price/cpi$value)*100

sonoma_log_returns=diff(log(sonoma$value))
sonoma_log_returns=c(0,sonoma_log_returns)
vallejo_log_returns=diff(log(vallejo$value))
vallejo_log_returns=c(0,vallejo_log_returns)

sonoma_real_returns=diff(log(sonoma_real))
sonoma_real_returns=c(0,sonoma_real_returns)
vallejo_real_returns=diff(log(vallejo_real))
vallejo_real_returns=c(0,vallejo_real_returns)

#converting returns to time series
sonoma_real_returns=ts(sonoma_real_returns, start=1978, end=2022, freq=4)
vallejo_real_returns=ts(vallejo_real_returns, start=1978, end=2022, freq=4)
```

Descriptive Statistics

```{r}
v.mean = mean(vallejo$value)
v.var = var(vallejo$value)
v.skew = skewness(vallejo$value)
v.kurt = kurtosis(vallejo$value)

s.mean = mean(sonoma$value)
s.var = var(sonoma$value)
s.skew = skewness(sonoma$value)
s.kurt = kurtosis(sonoma$value)

describe(vallejo$value)
```
```
````

````
```{r}
describe(sonoma$value)

```


```{r}
plot(sonoma$date,sonoma$value,type='l',main='House Prices in Solano and
````
````

```
Sonoma County',xlab='Year',ylab='Price (Q1 1995=100)',col='blue')
lines(vallejo$date,vallejo$value,type='l',main='House Prices in Solano County',xlab='Year',ylab='Price (Q1 1995=100)',col='red')
legend(c('bottomright'),inset=0.05,legend=c("Solano County", "Sonoma County"),
       col=c("red", "blue"), lty=1:1, cex=0.8)
```

```{r}
plot(vallejo$date,vallejo_log_returns,type='l',main='Log House Prices in Solano County and Sonoma County',xlab='Year',ylab='Price (Q1 1995=100)',col='red')
lines(sonoma$date,sonoma_log_returns,type='l',xlab='Year',ylab='Price (Q1 1995=100)',col='blue')
legend(c('bottomright'),inset=0.05,legend=c("Solano County", "Sonoma County"),
       col=c("red", "blue"), lty=1:1, cex=0.8)
```

ARIMA Models for Prices and Returns
```{r}
auto_sonoma_price <- auto.arima(sonoma_real)
auto_vallejo_price <- auto.arima(vallejo_real)
#auto_sonoma_returns <- auto.arima(sonoma_log_returns)
#auto_vallejo_returns <- auto.arima(vallejo_log_returns)
auto_sonoma_returns <- auto.arima(sonoma_real_returns)
auto_vallejo_returns <- auto.arima(vallejo_real_returns)
```

Seasonality Tests
```{r}
seasdum(sonoma_price, freq=4, autoarima = T)
seasdum(vallejo_price, freq=4, autoarima = T)
seasdum(sonoma_log_returns, freq=4, autoarima = T)
seasdum(vallejo_log_returns, freq=4, autoarima = T)
seasdum(sonoma_real_returns, freq=4, autoarima = T)
seasdum(vallejo_real_returns, freq=4, autoarima = T)
```

Fitting ARIMA Models
```{r}
```

````
auto_sonoma_price$coef
auto_vallejo_price$coef
auto_sonoma_returns$coef
auto_vallejo_returns$coef
```

## Examining Accuracy via RMSE
```{r}
#Checking Accuracy with RMSE
accuracy(auto_sonoma_price)[2]
accuracy(auto_vallejo_price)[2]
accuracy(auto_sonoma_returns)[2]
accuracy(auto_vallejo_returns)[2]
```

```{r}
auto_sonoma_price %>% forecast(h=20) %>% autoplot(xlab = "Time", ylab =
"Price", title = "Forecast of Sonoma Housing Prices (Inflation adjusted)
with ARIMA")
auto_sonoma_returns %>% forecast(h=20) %>% autoplot(xlab = "Time", ylab =
"Return")
```

```{r}
auto_vallejo_price %>% forecast(h=20) %>% autoplot(xlab = "Time", ylab =
"Price")
auto_vallejo_returns %>% forecast(h=20) %>% autoplot(xlab = "Time", ylab =
"Return")
```

## Cointegration
```{r}
coint1 = ca.jo(cbind(sonoma_price,vallejo_price),K=2)
coint2 = ca.jo(cbind(sonoma_log_returns,vallejo_log_returns),K=2)
summary(coint1)
summary(coint2)
```

Nominal housing prices are not cointegrated, but the log differenced prices
are.
````
````

```
## External Data Sources: Interest Rates, Approved Building Permits, and
Unemployment

```{r}
intrate = fredr(
  series_id = intrate_api,
  observation_start = as.Date(start),
  observation_end = as.Date(end)
)

unrate = fredr(
  series_id = unrate_api,
  observation_start = as.Date(start),
  observation_end = as.Date(end)
)

permits = fredr(
  series_id = permit_api,
  observation_start = as.Date(start),
  observation_end = as.Date(end)
)

gdp = fredr(
  series_id = gdp_api,
  observation_start = as.Date(start),
  observation_end = as.Date(end)
)


# Convert to quarterly to match our housing data
# Perform the slicing operation using row index
intrate <- intrate[which(seq_len(nrow(intrate)) %% 3 == 1), ]
unrate <- unrate[which(seq_len(nrow(unrate)) %% 3 == 1), ]
permits <- permits[which(seq_len(nrow(permits)) %% 3 == 1), ]



```

Testing Stationarity for External Data
```

````
```{r}
adf.test(intrate$value)
adf.test(unrate$value)
adf.test(permits$value)
adf.test(gdp$value)
```

We do not reject the null of non-stationarity for unrate and permits. So we
will need some transformations.

```{r}
intrate_t = diff((intrate$value))
unrate_t = diff((unrate$value))
permits_t = diff((permits$value))
gdp_t = diff(gdp$value)
```

## Testing Seasonality for External Data
```{r}
seasdum(intrate$value, freq=4, autoarima = T)
seasdum(unrate$value, freq=4, autoarima = T)
seasdum(permits$value, freq=4, autoarima = T)
seasdum(gdp$value, freq=4, autoarima = T)

seasdum(intrate_t, freq=4, autoarima = T)
seasdum(unrate_t, freq=4, autoarima = T)
seasdum(permits_t, freq=4, autoarima = T)
seasdum(gdp_t, freq=4, autoarima = T)
```

Nominal interest rate may have seasonality. First integration interest rate
however does not.

First integration proves stationary for all variables.

## Creating ARIMA Models with Different Exogenous Variables
```{r}
auto_sonoma_exog = auto.arima(sonoma_price,
xreg=cbind(intrate$value,unrate$value,permits$value,gdp$value))
auto_sonoma_exog$coef
auto_vallejo_exog = auto.arima(vallejo_price,
````
````

```
xreg=cbind(intrate$value,unrate$value,permits$value,gdp$value))
auto_vallejo_exog$coef
```

For Sonoma, as we might expect, the ARIMA models show a negative coefficient on interest rate and unemployment rate, but a positive (though slight) positive coefficient on building permits.
Solano interestingly shows a positive coefficient on unemployment, and a negative coefficient on permits.

Models Using Stationary Time Series

```{r}
auto_sonoma_ex2 = auto.arima(sonoma_log_returns[2:180],
xreg=cbind(intrate$value[2:180],unrate_t,permits_t,gdp_t))
auto_sonoma_ex2$coef

auto_vallejo_ex2 = auto.arima(vallejo_log_returns[2:180],
xreg=cbind(intrate$value[2:180],unrate_t,permits_t,gdp_t))
auto_vallejo_ex2$coef
```

Cointegration Tests for Nominal and Returns

```{r}
coint1 = ca.jo(cbind(sonoma_price,vallejo_price),K=2)
coint2 = ca.jo(cbind(sonoma_log_returns,vallejo_log_returns),K=2)
summary(coint1)
summary(coint2)
```

VAR Model for Nominal

```{r}
data <- data.frame(sonoma = diff(sonoma_price), vallejo =
diff(vallejo_price))
# 12 Max Lags
lag_opt <- VARselect(data, lag.max = 12, type = "both")$selection[1]
# Fit the VAR model
var_model <- VAR(data, p = lag_opt, type = "both")
summary(var_model)
```

````
## Granger Causality for Nominal
```{r}
granger_test_sonoma <- causality(var_model, cause = "vallejo")
print(granger_test_sonoma)
```
```{r}
granger_test_vallejo <- causality(var_model, cause = "sonoma")
print(granger_test_vallejo)
```
## VAR Model for Returns
```{r}
data <- data.frame(sonoma = sonoma_log_returns, vallejo = vallejo_price)
# 12 Max Lags
lag_opt <- VARselect(data, lag.max = 12, type = "both")$selection[1]
# Fit the VAR model
var_model <- VAR(data, p = lag_opt, type = "both")
summary(var_model)
```

## Granger Causality for Returns
```{r}
granger_test_sonoma <- causality(var_model, cause = "vallejo")
print(granger_test_sonoma)
```
```{r}
granger_test_vallejo <- causality(var_model, cause = "sonoma")
print(granger_test_vallejo)
```

## GARCH Models
```{r}
g1_sonomareal <- garchFit(~garch(1,1),data=sonoma_real_returns,trace=F)
g1_vallejoreal <- garchFit(~garch(1,1),data=vallejo_real_returns,trace=F)
plot(g1_sonomareal, which=1)
plot(g1_vallejoreal, which=1)
```
```{r}
g1_sonomareal@fit$coef
g1_sonomareal@fit$ics
g1_vallejoreal@fit$coef
````
````

````
g1_vallejoreal@fit$ics
```
```{r}
gres1=residuals(g1_sonomareal,standardize=T)
gres2=residuals(g1_vallejoreal,standardize=T)
hist(gres1)
hist(gres2)
print("Sonoma (1,1) Residuals Skewness and Kurtosis")
skewness(gres1)
kurtosis(gres2)
print("Solano (1,1) Residuals Skewness and Kurtosis")
skewness(gres1)
kurtosis(gres2)
```
## Ljung Box Test
```{r}
Box.test(gres1,type='Ljung')
Box.test(gres2,lag=10,type='Ljung')
Box.test(gres1^2,lag=10,type='Ljung')
Box.test(gres2^2,lag=10,type='Ljung')
```
## IGARCH
```{r}
#IGARCH
#source("Igarch.R")
#i2_sonomareal <- Igarch(sonoma_real_returns)

spec = ugarchspec(variance.model=list(model="iGARCH", garchOrder=c(1,1)),
mean.model=list(armaOrder=c(1,1)))

i1_sonomareal <- ugarchfit(spec, data = sonoma_real_returns)
i1_sonomareal

resi=i1_sonomareal@fit$residuals/i1_sonomareal@fit$sigma
print('iGARCH Sonoma Standardized Residuals Skewness and Kurtosis')
skewness(resi)
kurtosis(resi)

plot(i1_sonomareal, which=8)
```
```{r}
````

```
i1_vallejoreal <- ugarchfit(spec, data = vallejo_real_returns)
i1_vallejoreal

resi1=i1_vallejoreal@fit$residuals/i1_vallejoreal@fit$sigma
print('iGARCH Solano Standardized Residuals Skewness and Kurtosis')
skewness(resi1)
kurtosis(resi1)

plot(i1_vallejoreal, which=8)
```
TGARCH
```{r}
#TGARCH
#source("Tgarch11.R")
#t1_sonomareal <- Tgarch11(sonoma_real_returns)

spec2 = ugarchspec(variance.model=list(model="fGARCH",submodel="TGARCH",
garchOrder=c(1,1)), mean.model=list(armaOrder=c(1,1)))

t1_sonomareal <- ugarchfit(spec2, data = sonoma_real_returns)
t1_sonomareal

resi2=t1_sonomareal@fit$residuals/t1_sonomareal@fit$sigma
skewness(resi2)
kurtosis(resi2)

plot(t1_sonomareal, which=8)
```
```{r}
t1_vallejoreal <- ugarchfit(spec2, data = vallejo_real_returns)
t1_vallejoreal

resi3=t1_vallejoreal@fit$residuals/t1_vallejoreal@fit$sigma

skewness(resi3)
kurtosis(resi3)

plot(t1_vallejoreal, which=8)
```
```{r}
plot(i1_sonomareal, which=3)
```

````
plot(i1_vallejoreal, which=3)
```
```{r}
p1_sonoma <- ugarchforecast(fitORspec=t1_sonomareal,n.ahead=10)
p1_vallejo <- ugarchforecast(fitORspec=t1_vallejoreal,n.ahead=10)
```


```{r}
plot(p1_sonoma, which=1)
plot(p1_vallejo, which=1)
plot(p1_sonoma, which=3)
plot(p1_vallejo, which=3)
```
## Value-at-Risk Analysis

### Sonoma County VaR and ES

### Risk Metrics Method
```{r}
source("RMfit.R")
sonoma_rm=RMfit(sonoma_real_returns)
```
## Trying with iGARCH

```{r}
spec = ugarchspec(variance.model=list(model="iGARCH", garchOrder=c(1,1)),
mean.model=list(armaOrder=c(1,1)))

i1_sonomareal <- ugarchfit(spec, data = sonoma_real_returns)
i1_vallejoreal <- ugarchfit(spec, data = vallejo_real_returns)

p1_sonoma <- ugarchforecast(fitORspec=i1_sonomareal,n.ahead=10)
p1_vallejo <- ugarchforecast(fitORspec=i1_vallejoreal,n.ahead=10)

p1_sonoma@forecast$seriesFor
p1_sonoma@forecast$sigmaFor

source("RMeasure.R")
sonoma_rmg=(RMeasure(p1_sonoma@forecast$seriesFor[1],
p1_sonoma@forecast$sigmaFor[1]))
```
````
````

```
### Empirical Quantile
```{r}
sonoma_quant = quantile(sonoma_real_returns,c(0.95,0.99,0.999))
sonoma_quant
```
```{r}
sonoma_med = 829000
sonoma_maxvar = max(sonoma_quant[1],sonoma_rmg$results[1,2],0.03257497)
sonoma_maxes = max(sonoma_rmg$results[1,3],0.04085033)
print(paste0("Max VaR for Sonoma: ",round(sonoma_maxvar,5)*100,"%"))
print(paste0("Max ES for Sonoma: ",round(sonoma_maxes,5)*100,"%"))
sonoma_maxvar*sonoma_med
sonoma_maxes*sonoma_med
```
Solano County VaR and ES

Risk Metrics Method
```{r}
source("RMfit.R")
# RMfit(vallejo_real_returns)
```
Econometric (iGARCH) Modeling

```{r}
p1_vallejo@forecast$seriesFor
p1_vallejo@forecast$sigmaFor
source("RMeasure.R")
vallejo_rmg=RMeasure(p1_vallejo@forecast$seriesFor[1],
p1_vallejo@forecast$sigmaFor[1])
```
Empirical Quantile
```{r}
vallejo_quant=quantile(vallejo_real_returns,c(0.95,0.99,0.999))
vallejo_quant
```
```{r}
vallejo_med = 570922
vallejo_maxvar = max(vallejo_quant[1],vallejo_rmg$results[1,2],0.04512308)
vallejo_maxes = max(vallejo_rmg$results[1,3],0.05658616)
print(paste0("Max VaR for vallejo: ",round(vallejo_maxvar,5)*100,"%"))
print(paste0("Max ES for vallejo: ",round(vallejo_maxes,5)*100,"%"))
```
```

````
vallejo_maxvar*vallejo_med
vallejo_maxes*vallejo_med
```
## Conclusion and Managerial Implications

Returning to our ARIMA models for forecasting returns:

```{r}
auto_sonoma_price <- auto.arima(sonoma_real)
auto_vallejo_price <- auto.arima(vallejo_real)
auto_sonoma_returns <- auto.arima(sonoma_real_returns)
auto_vallejo_returns <- auto.arima(vallejo_real_returns)
```

```{r}
sonoma_price_forecast <- forecast(auto_sonoma_price,10)
sonoma_returns_forecast <- forecast(auto_sonoma_returns,10)
vallejo_price_forecast <- forecast(auto_vallejo_price,10)
vallejo_returns_forecast <- forecast(auto_vallejo_returns,10)

sonoma_returns_forecast
vallejo_returns_forecast
```
````
````