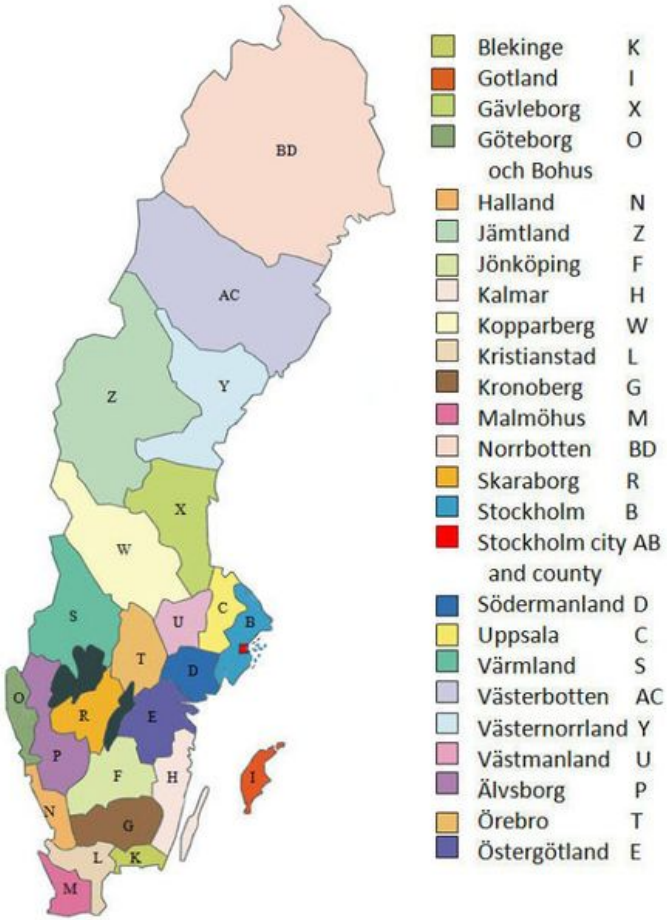
KTH ROYAL INSTITUTE
OF TECHNOLOGY

Construction of daily hedonic housing indexes for apartments in Sweden

Mo Zheng
Division of Building and Real Estate Economics
School of Architecture and the Built Environment
Royal Institute of Technology, Sweden
mo.zheng@abe.kth.se

Introduction



- Sweden is the third-largest country in EU by area with a population of 10 million
- 2.2 million people live in Stockholm.
- Interest Rate: - 0.5%
- Mortgage Rate:

1996 over 10%, 2004-2008 5%, Now -0.5%

- Inflation Rate: 1.7%
- Loan Tenure:

140 years DROPPED to 105 years

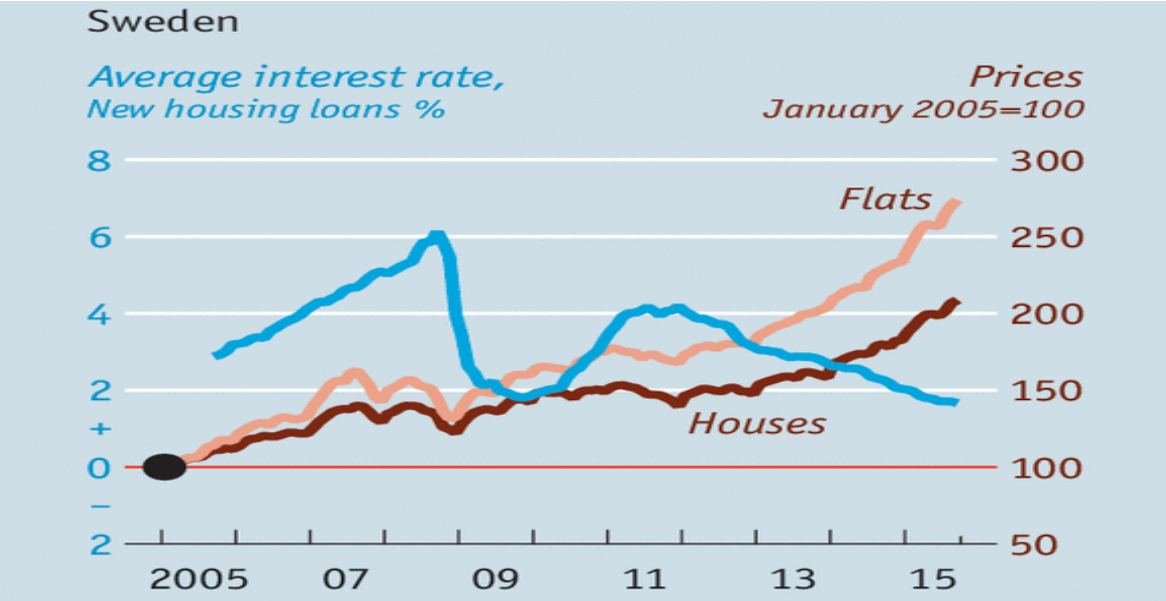
- Loan Amount: Average 5-6 times of annual income

For example:

KTH, located inner Stockholm Area,

120,00 Euro per square meters

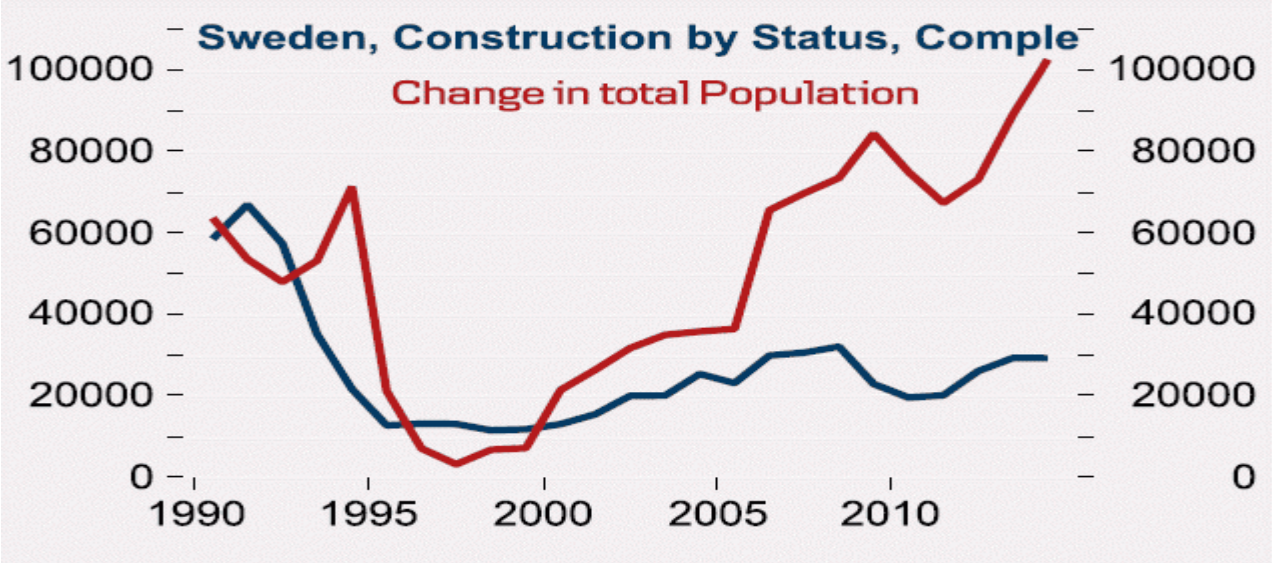
Introduction



Source from: Statistics Sweden

Interest rate is historical low and the home loan rates was more than 10% in 1996, dropped to less than 5% during 2004-2008 and reached to less than 2% now according to Statistics Sweden.

Introduction



Source from: Statistics Sweden

Over regulated Rental Market
Minimum waiting time: 8.5 years

Current House Price Index

- The Statistics Sweden produces quarterly real estate price index for buildings for seasonal and secondary use, one and two dwelling buildings for permanent living by region back to 1986 first quarter based on tax authority only with no further details and control over different characteristics of the property.

- Another house price index is the monthly updated price index for condominiums commonly referred by media, and published by Svensk Mäklarstatistik, an organization owned by Swedish real estate agencies and institutions but this index only presents mean and median prices.

- Nasdaq OMX Valueguard-KTH Housing Index (HOX) Sverige

Methodology

According to the literatures, the hedonic price equation applying time dummy variable method could be written as follows:

$$ln(p_{i,t}) = \beta_0 + \sum_{k=1}^{K} \beta_k Z_{i,k,t} + \sum_{t=1}^{T} \delta_\tau TD_{i,\tau} + \varepsilon_{i,t}$$

$i=1,\ldots.N,\ t=0,\ldots..T$

Where $ln(p_{i,t})$ is the log form of price as dependent variable. β_0 is the intercept term, K is the exogenous explanatory characteristics, and $Z_{i,k,t}$, k=1,…K as a set of quality-variables which could be continuous variables such as area and a set of dummy variables which is category as 0 and 1.

Time dummy variable method

$TD_{i\tau}$ is the time-dummy variables , which measures the effect of time (daily changes in this paper) on dependent variable, starting from the base period t=0, the coefficient of time dummy parameters shows the change over time. Kennedy's formula of interpreting the dummy variables in semi-log form is applied, in other word, from t=$_{T_m}$ to t=$_{T_n}$, we calculated the price change as follows:

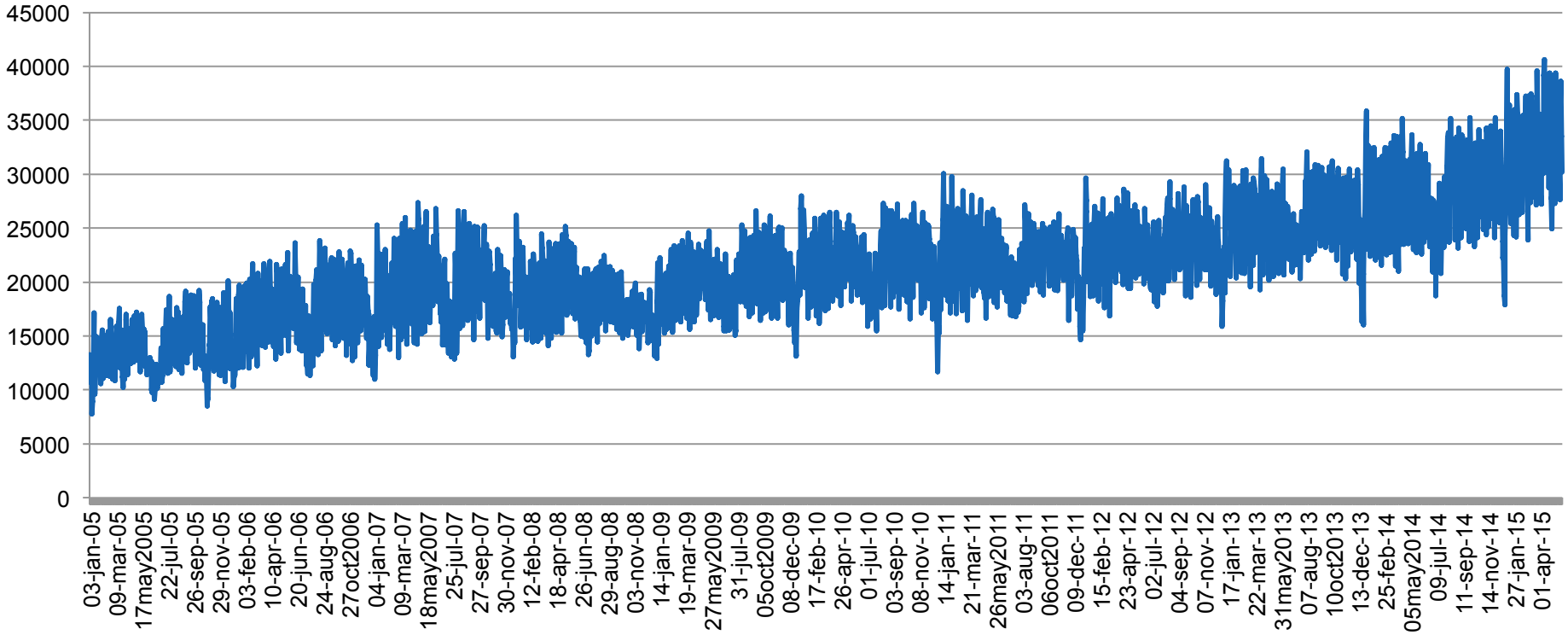
$$P_{TD_{m-n}} = p_m(z)/p_n(z) = \exp(\delta_m - \delta_n)$$

Where $_z$ is the quality configuration, the estimated price index for $_{T_m}$ relative to $_{T_n}$ is the exponential of the difference of coefficient of time-dummy variable.

Data

- We use a unique cross-sectional time series dataset with 593,930 observations of all arms-length transactions of apartments in Sweden from January 2005 to May 2015.
- We take a naïve data cleaning process by deleting top and bottom 1% of the transaction price
- The lowest : 95,000 SEK / and the most expensive: 6,2million SEK. Average: 1.5 million SEK and median 1.25 million SEK.
- In average, there are 227 transactions per day in our dataset to construct a daily housing index.
- Average size of the apartment is 65 m^2 , the smallest unit is 24m^2 , and the largest one is 138m^2 .

Data



Average daily transacted price per square meter (in SEK)

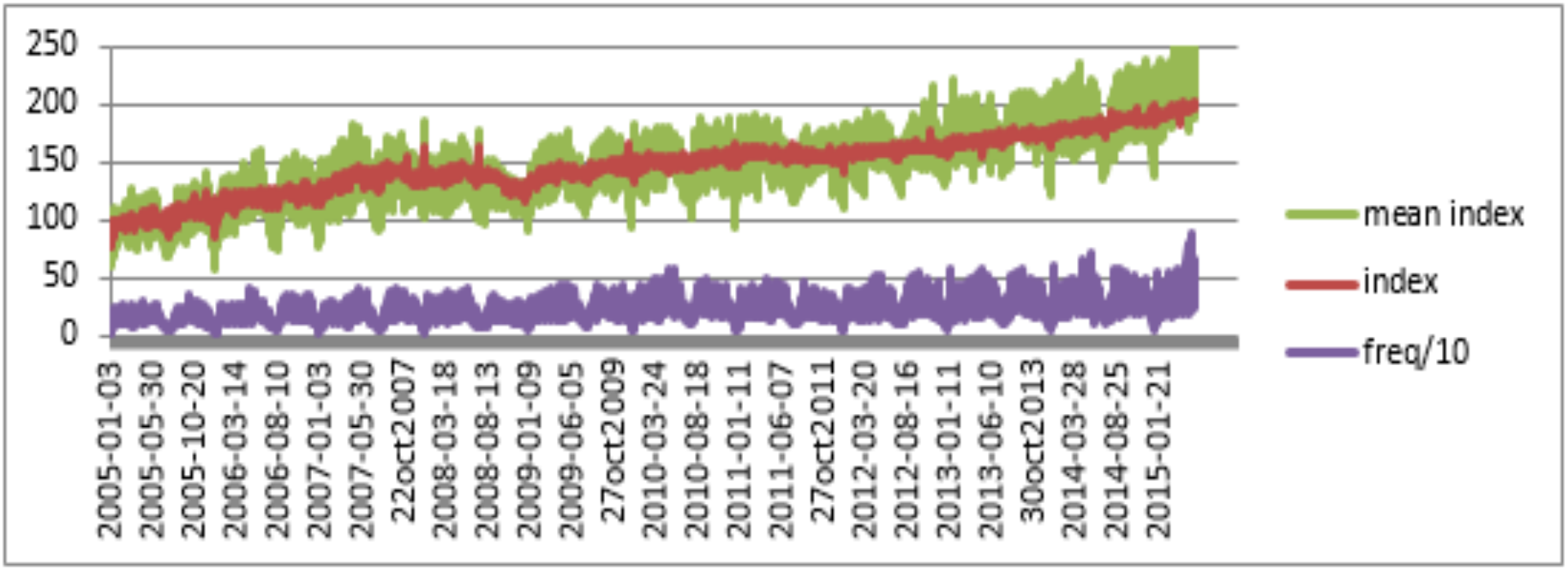
Data

Table 1. Descriptive statistics

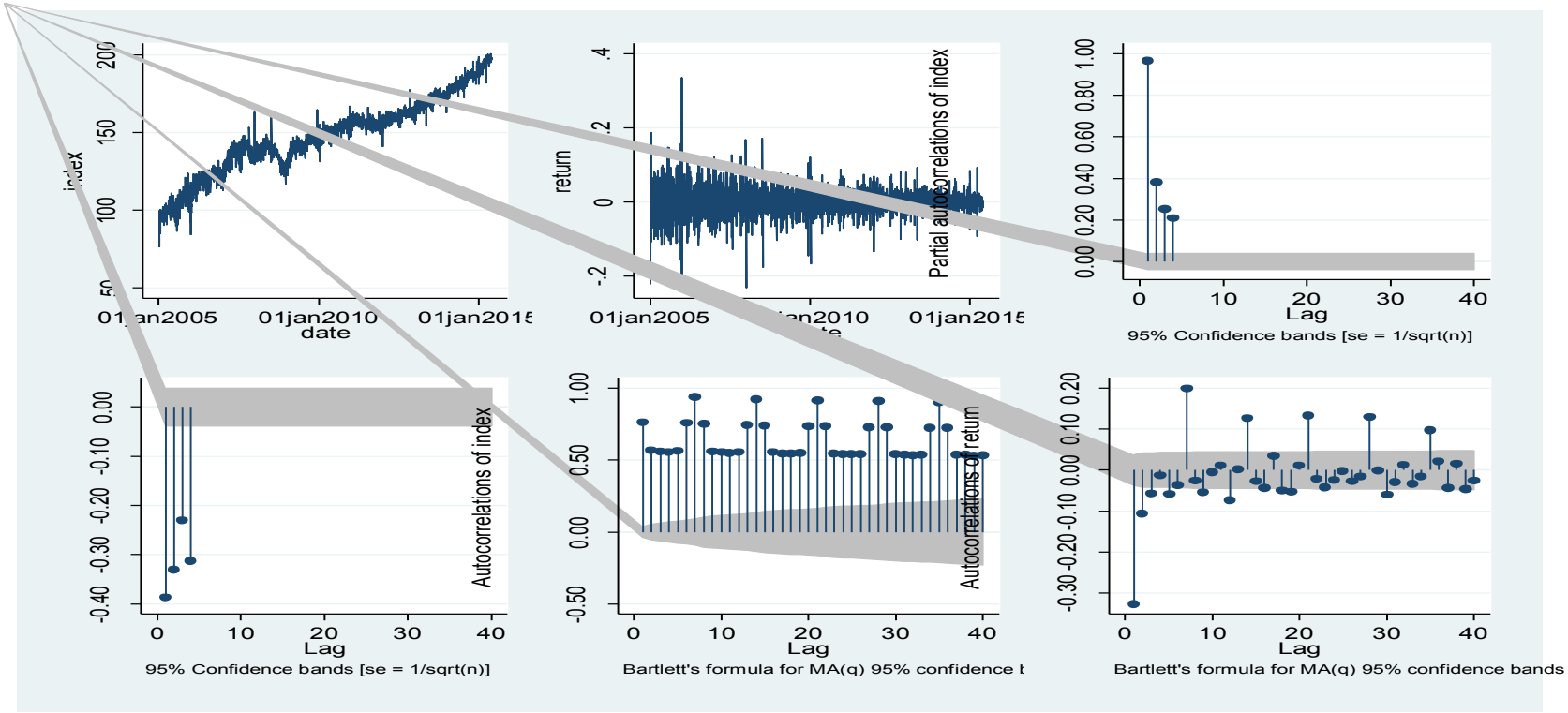
Variable	description	Mean	Std. Dev.	Min	Max
lnprice	logrithim of price	13.93	0.78	11.46	15.65
lnarea	logrithim of living area	4.15	0.35	3.18	4.93
lnmfee	logrithim of monthly fee	8.10	0.37	7.00	8.88
Room	Number of rooms	2.49	0.99	1.00	6.00
Age	Age	44.88	25.69	0	115
X	longtitute	6524257	203468	6137782	7595126
seaview	Dummy: Seaview	0.00	0.03	0	1
balcony	Dummy: Balcony	0.14	0.35	0	1
elevator	Dummy: Elevator	0.40	0.49	0	1
top	Dummy: top floor	0.29	0.45	0	1
basement	Dummy: basement	0.00	0.03	0	1
first	Dummy: ground floor	0.22	0.41	0	1
WW 2	Dummy: 1900-1939	0.16	0.37	0	1
Recover	Dummy: 1940-1959	0.26	0.44	0	1
MillionPro	Dummy: 1960-1975	0.25	0.43	0	1
HighSub	Dummy: 1976-1990	0.13	0.34	0	1
nosub	Dummy: apartment built after 1990	0.20	0.40	0	1
kommun	Dummy: cummunity code	867	735	114	2584

Result

We run the hedonic regression on our dataset, and Adjusted-R^2 is around 80%, which means 80% of the price determinates could be explained by the model.

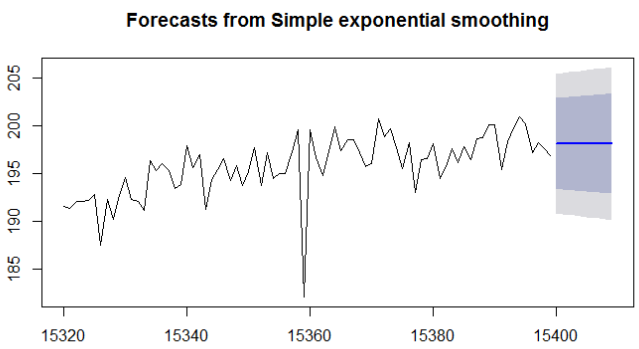


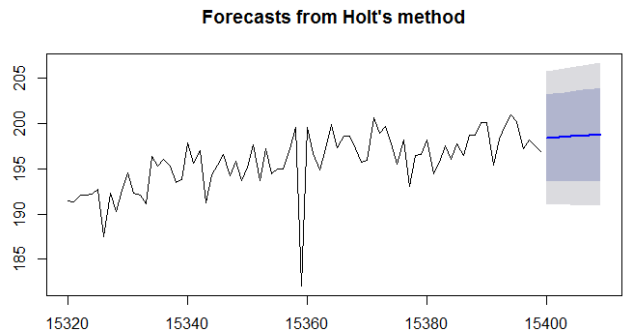
Forecasting

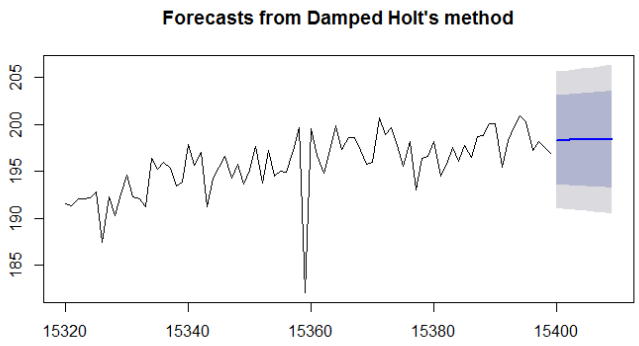


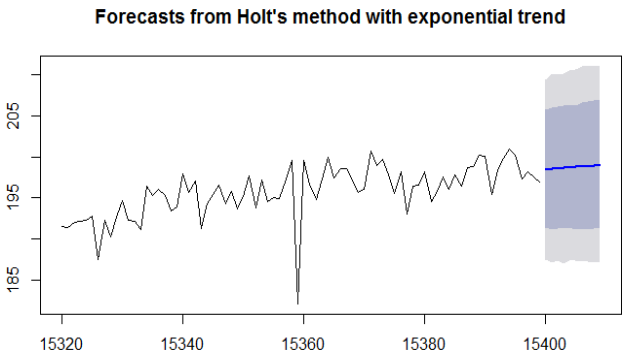
Non- Stationary time series, take the transformation of differencing, and generate new series : Daily index Return. ACF and PACF performed before and after the transformation.

Forecasting

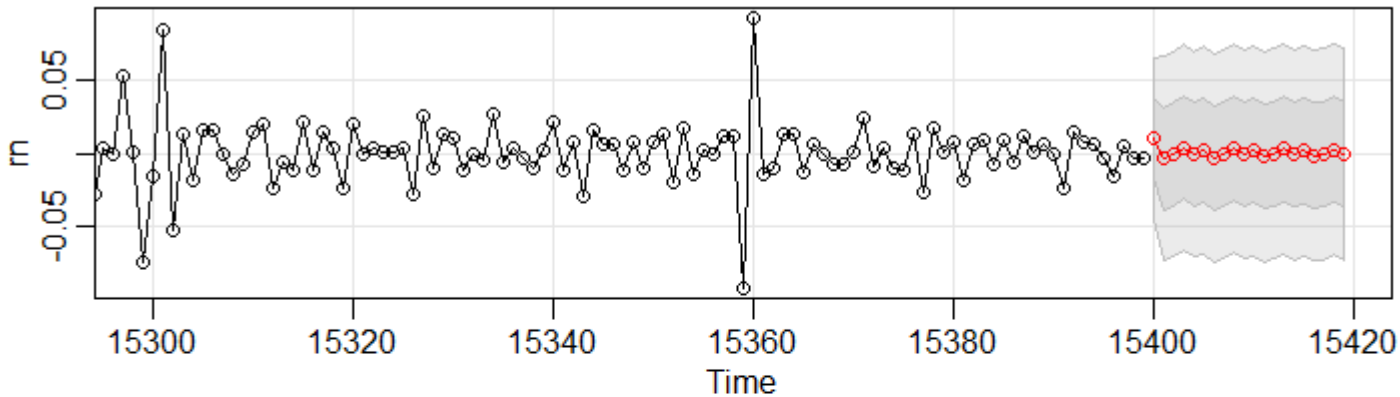








Forecasting by ARIMA Model



AIC and BIC statistics show that ARIMA model fits best

Forecast return for 20 days ahead based on an ARIMA(5,0,5)

Further research Questions

Obtain new transaction records from 2015 May till now into the dataset and perform time dummy hedonic regression to generate the latest housing index 2.0 and compare the result with the predicted values.

Perform VAR model on the daily index together with other financial assets such as interest rate, mortgage rate, related stocks or bonds and exchange rate with USD and Euro.

Thank you!