Understanding the Impact of Housing Starts in British Columbia

Lucas Mayall & Christian Koutsandraes

2022/12/08

Introduction

This report on housing starts in British Columbia centers around analysis of a time series that tracks the number of new residential construction projects in the province of British Columbia, Canada, at quarterly intervals. This time series is of particular interest because it allows examination of the relationship between housing starts and rent costs in British Columbia. By looking at the data over the 67-year period from 1955 to 2022, one can observe how changes in the number of housing starts have impacted the average rent prices in the province.

Housing starts data is analyzed as opposed to housing completion data because housing starts provide a more real-time picture of the state of the housing market. Housing starts represent the number of new residential construction projects that are being initiated, which allows tracking of the current demand for housing and the overall health of the housing market in British Columbia. In contrast, housing completion data only provides information about the number of projects that have been completed, which may not directly reflect the current state of the housing market as construction takes time. Overall, housing starts are a more useful and relevant indicator of the current state and future trends of the housing market in British Columbia.

Importance

The importance of this time series lies in its ability to provide valuable insights into the housing market in British Columbia. The high housing costs and rising cost of living in the province are major concerns for many people[1], and this data allows us to better understand the factors that drive these costs. Notably, this information can be useful for anyone interested in buying a home, as well as any industries with an overarching focus on real estate such as agents, developers, and policy makers[2]. Further, students are particularly vulnerable to the rising housing costs due to their often low incomes and limited financial resources[3]. This time series data can provide valuable insights into the housing market in British Columbia and help us to better understand the factors driving these costs, allowing us to make more informed decisions about housing supply and demands.

Focus

The analysis will focus on multi-family starts for several reasons. Firstly, they are gaining in popularity due to the increasing demand for dense, urban living[4]. Secondly, they are the most space-effective way to add new housing[4], which is important in a densely populated area like British Columbia. Finally, multi-family starts are considered the future of urban development[5], as they allow for more efficient use of land and resources.

Multi-family starts are broken down into three categories:

- Apartments: multi-unit buildings where each unit is individually owned
- Row houses: attached houses that share a common wall with their neighbors
- Semi-detached houses: attached houses with their own separate entrance

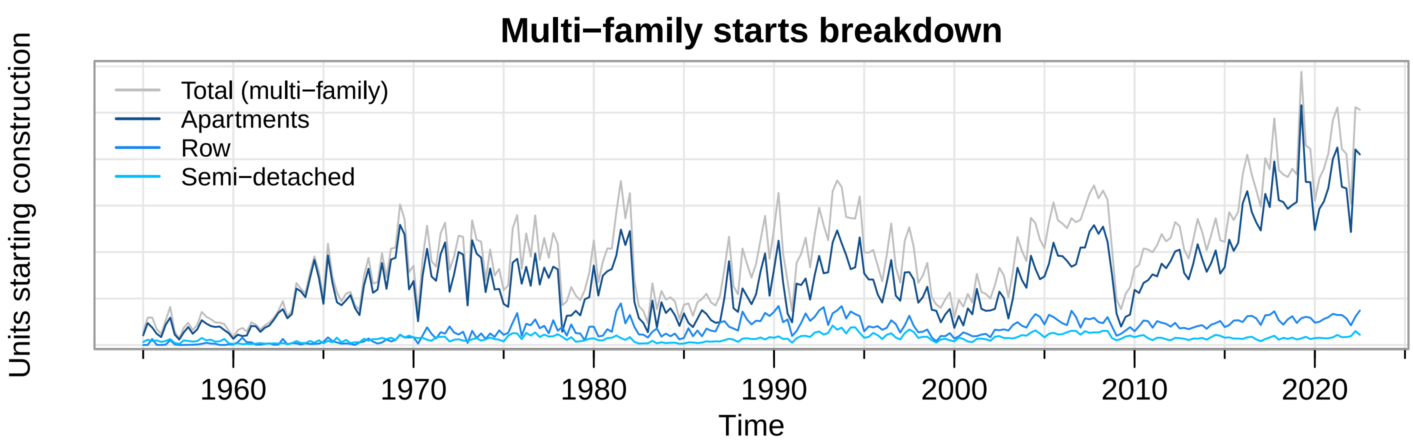


Figure 1: Breakdown of multi-family starts

As shown in figure 1, apartment starts have increased rapidly since 2000, though trends for row and semi-detached houses are difficult to discern. However, it is clear that these two types of units are far less popular than apartments, and their popularity does not appear to have increased relative to combined demands over time. The trends depicted in Figure 1 highlight the growing demand for multi-family housing in British Columbia due dominantly to apartments.

Finally, an optimal SARIMA model will be found for multi-family housing starts. The data will be transformed, and ACF and PACF will be used to assess model diagnostics and compare models. Upon reaching an optimal model, it will be used to predict the next 10 quarters of the time series.

Time series overview

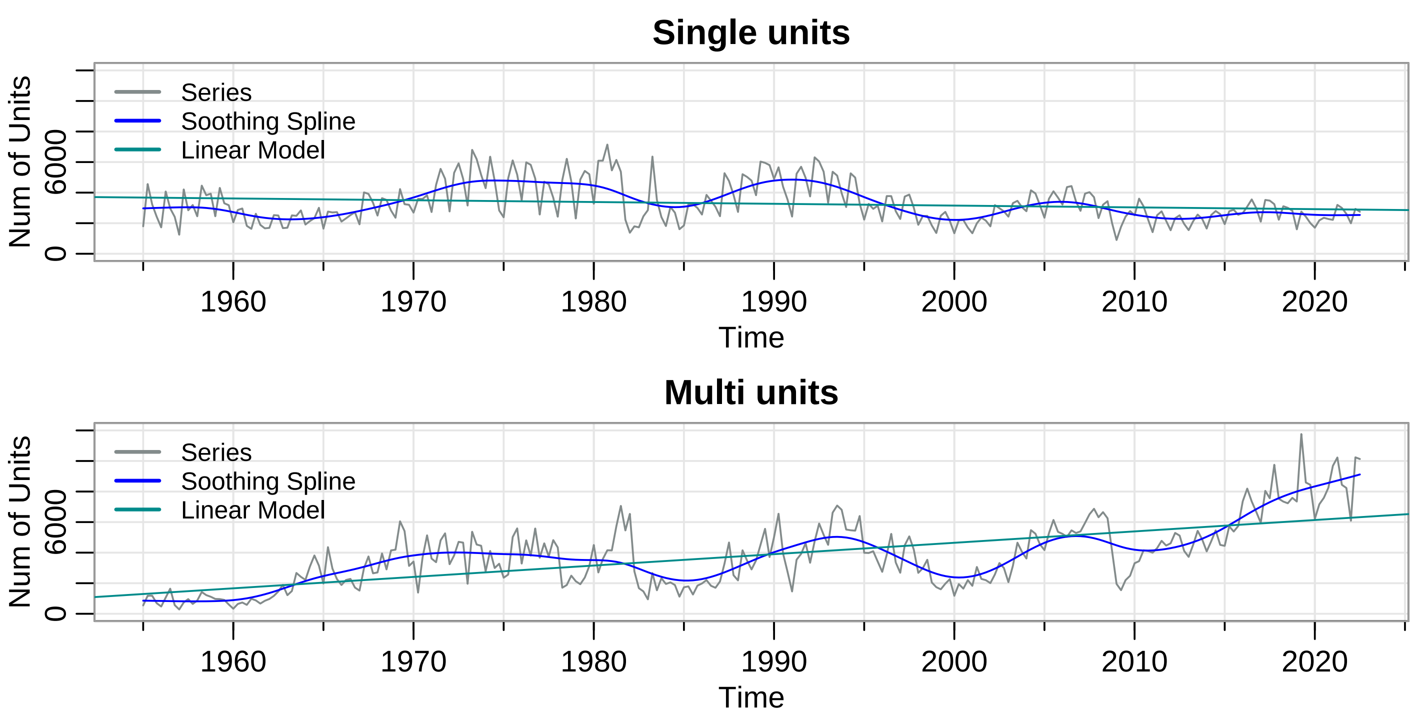


Figure 2: Time series of focus

The Figure 2 plots of single-family and multi-family housing starts in British Columbia provide insight into the changing preferences and needs of consumers in the province. The first plot shows that the number of single-family housing starts began to decline after the year 2000, while the second plot shows that the number of multi-family housing starts began to increase during the same period. These assumptions are further supported by the slope of the linear models respectively. These trends suggest that there has been a shift in the types of homes that consumers are looking for which has influenced the decisions of developers in the housing market.

One possible explanation for these trends is the impact of the economic downturn in the year 2000. The downturn may have caused a decrease in demand for single-family homes, as consumers may have become more cautious about making large financial commitments such as buying a house. In essence, the downturn definitively marks a shift towards more affordable housing options, such as multi-family homes, as people presumably looked for ways to save money on housing costs. As a result, developers have responded to these changing consumer needs by focusing more on building multi-family homes and less on single-family homes.

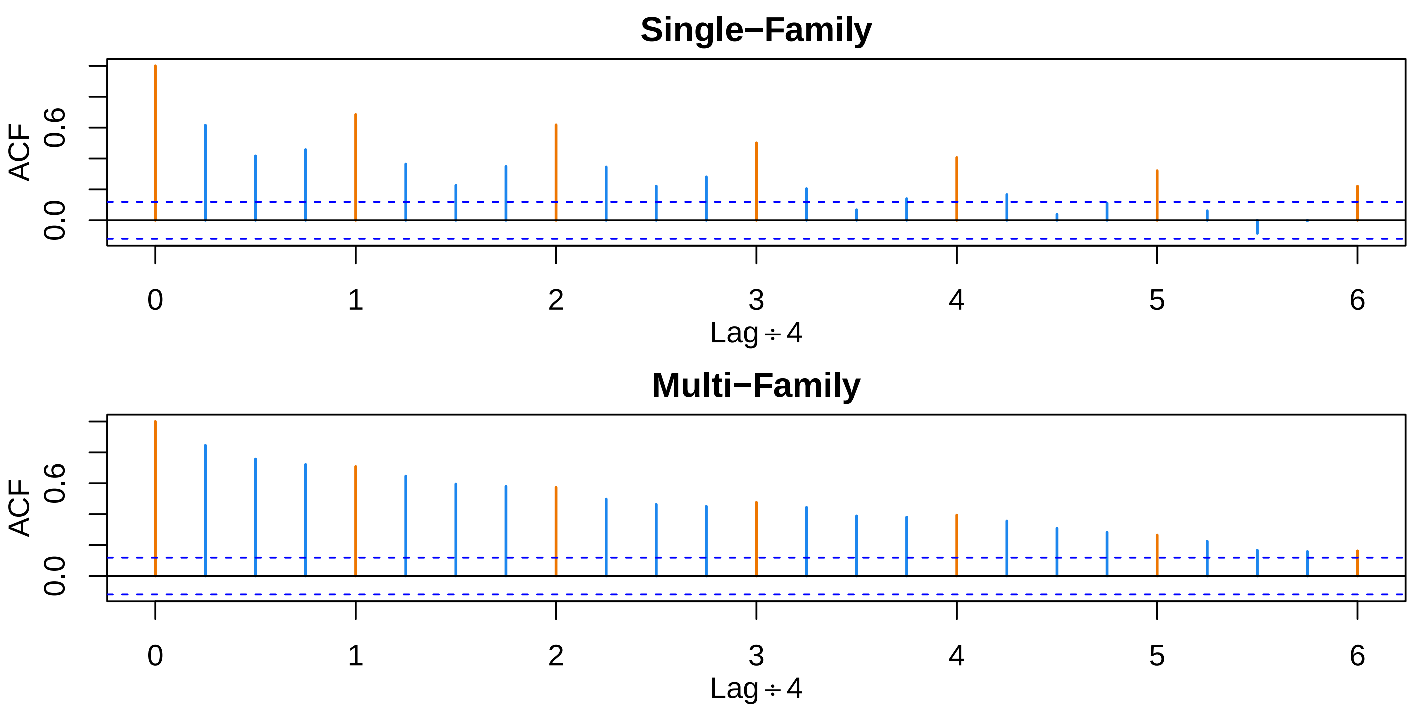


Figure 3: Time series ACF

The first plot in Figure 3 shows the ACF of single-family home starts, and it reveals a clear seasonality in the construction of these homes. Specifically, the ACF indicates that the number of starts tends to peak at 4 quarter (yearly) intervals. The second plot shows the ACF of multi-family construction starts, and it reveals only slight seasonality. This suggests that the construction of multi-family homes is not as strongly influenced by seasonal factors as single-family homes. This pattern likely reflects the weather conditions in the region, which are more conducive to breaking ground during the warmer months[6]. However as multi-family constructions can take several years, there is less incentive to avoid certain seasons as working through them is inevitable. It is also possible that the construction of multi-family homes is more subject to market demand and other economic factors, rather than being primarily influenced by the weather.

Bringing in other data

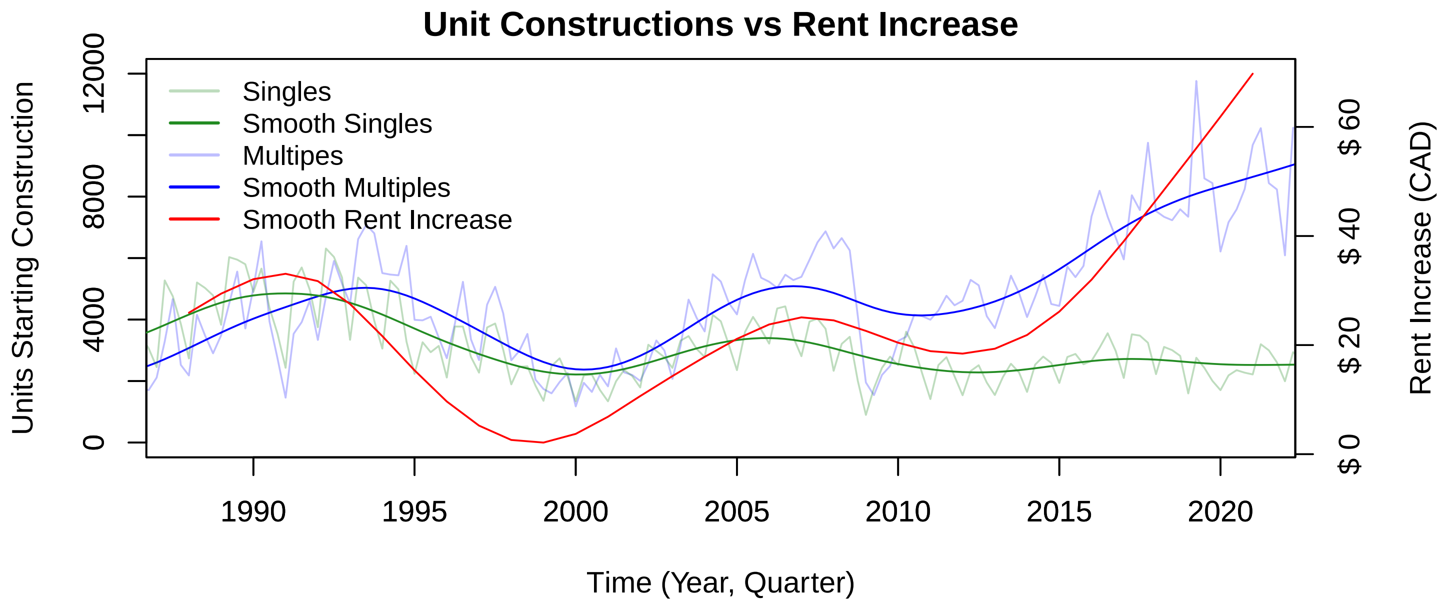


Figure 4: Constructions relationship with rent

Figure 4 shows the relationship between housing starts and quarterly rent changes in British Columbia. The plot reveals that multi-family housing starts and rent prices appear to be more correlated than single-family housing starts, which are only slightly correlated. This could be due to the fact that units in multi-family constructions are more often rentals versus fully detached single-family homes, which are commonly owned by the inhabitant[7]. Despite the different levels of correlation, both housing starts and rent prices appear to be correlated due to their local maximum and minimum values being offset by, at most, 4 years. This shows that they are correlated and likely caused by similar factors, despite feeling the effects of those factors at different rates.

Moreover, Figure 5 shows the relationship between housing starts and inflation in British Columbia. The plot appears to reveal that the correlation between these two variables is much weaker, or non-existent, when compared to the correlation between housing starts and rent, as shown in Figure 4. However, they are both likely influenced by similar economic factors, despite being mostly uncorrelated.

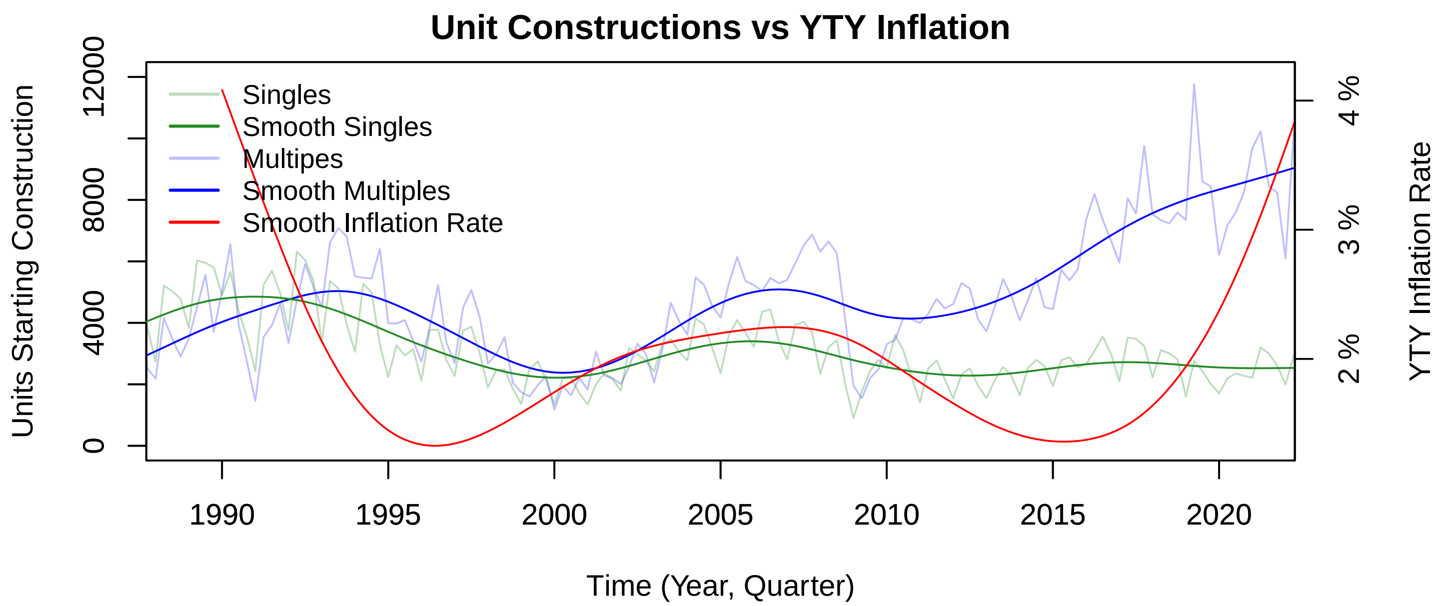
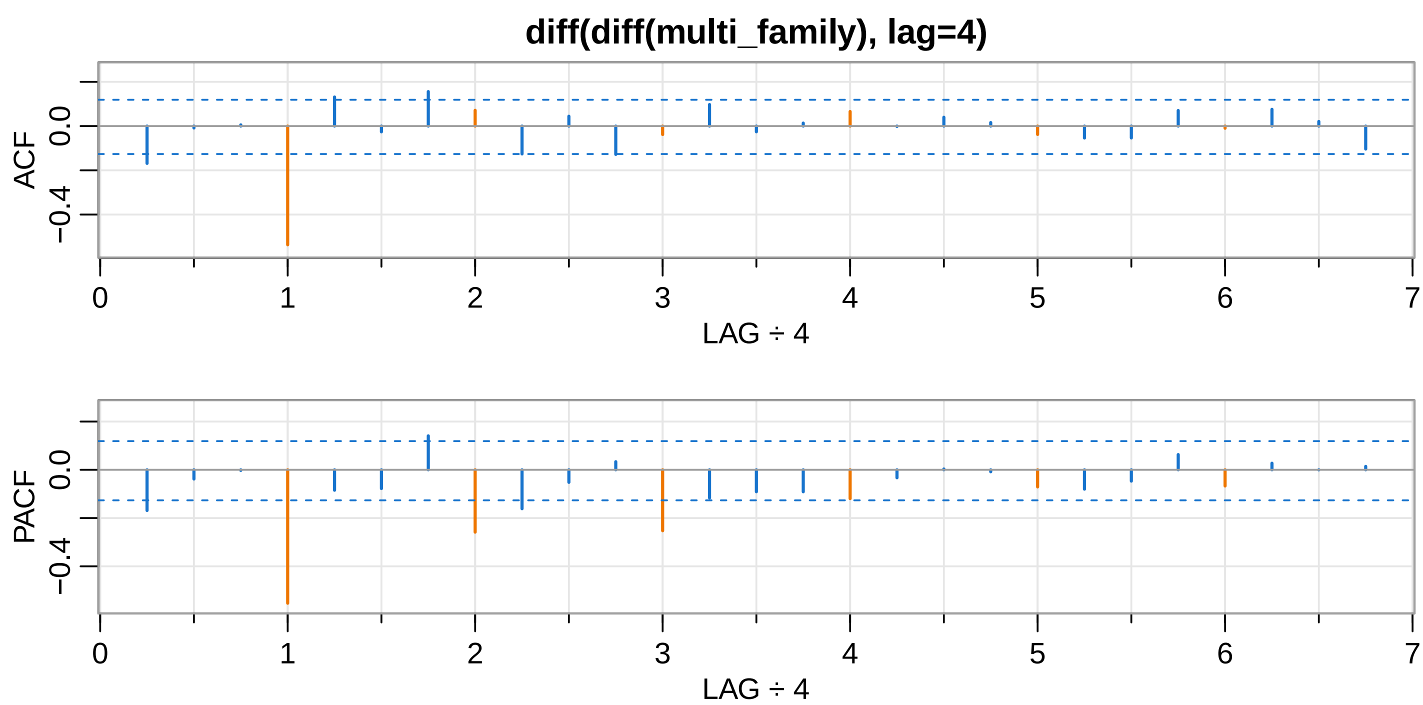


Figure 5: Constructions relationship with inflation

Transformations



To determine an adequate SARIMA model for a time series, the 1st difference is taken to de-trend the data. The seasonal difference is then taken at $S = 4$ to remove the remaining seasonal component. This helps identify the appropriate values for the d and D parameters. The ACF and PACF plots of the differenced time series are then analyzed to attempt determine the appropriate values for the remaining parameters in the SARIMA model.

Here are the key findings from the analysis of the ACF and PACF plots, referring to the fully transformed data unless otherwise specified:

ACF:

- The ACF of the first difference still shows a large seasonality at the seasonal points, however it does remove a majority of the correlation, suggesting a value of $d = 1$.
- The ACF shows the seasonal difference did not remove a massive amount of correlation, suggesting a value of $D = 0$ or $D = 1$ for the SARIMA model. From testing, $D = 0$ was found to be optimal.
- The ACF appears relatively stationary except for at the first seasonal period where there is significant negative correlation, suggesting a value of $P = 1$ for the SARIMA model.
- The in-season portion of the ACF shows the correlation cutting off at a lag of 1, suggesting a value of $p = 1$ for the SARIMA model.

PACF:

- The PACF shows significant negative correlation at the seasonal periods, specifically the first period, but it tails off.
- The in-season portion of the PACF shows the correlation cutting off at a lag of 1 or 2, suggesting potential values of $P = 1$ and $Q = 2$ for the SARIMA model.

Based on this analysis and testing, a SARIMA$(1, 1, 2)$ $(1, 0, 2)_{12}$ model was chosen.

Fit the model

In order to find the most accurate and reliable model, two separate techniques were used. The first, an analytical approach, involved starting with the model discussed in the previous section, performing diagnostics, and investigating alternative models. The second, a data-driven approach, involved iterating over all possible SARIMA models (with certain restrictions) and limiting the order of each parameter to a maximum of 3.

Analytical Approach

After testing the initial SARIMA$(1,1,2)$ $(1,0,2)_4$ model it was found to be adequate, however the p-values for the MA(2) and SMA(2) terms were not significant (p-value > 0.05). As a result, these terms were removed from the model, resulting in the revised model, SARIMA$(1,1,1)$ $(1,0,1)_4$. All terms in this model were found to be very significant with the p-values ≈ 0, change in information criteria was $< 0.1\%$, and MSE increased by $< 2\%$. In addition, the residuals of the revised model were found to be normally distributed and uncorrelated. The minor changes to the models performance, the residual analysis, and the lack of overfitting indicate that the model is a good fit for the data and is the optimal model given this approach.

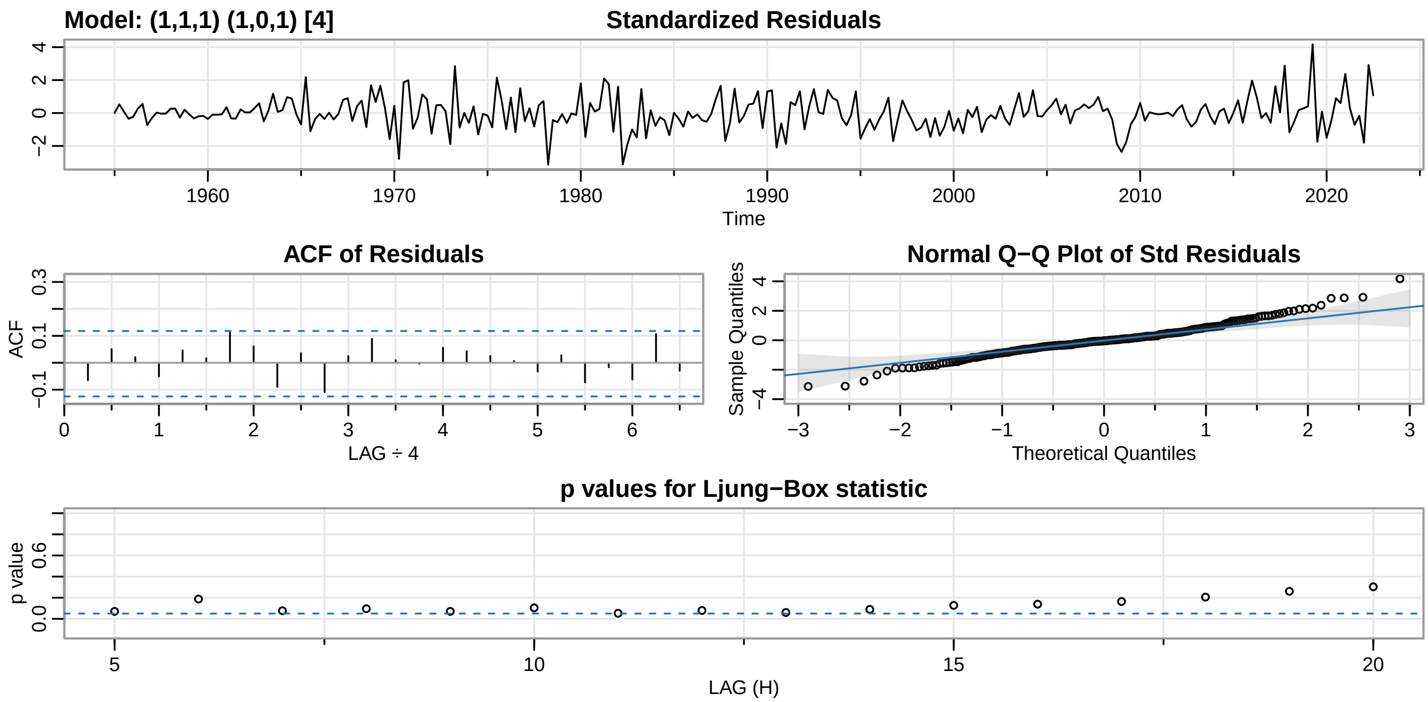


Figure 6: Final model details

Data driven approach

For the data-driven approach many different models were generated by iterating over all possible models where the order of each parameter was at most 3, giving 1024 unique models. These models were compared based on their AIC, AICc, and BIC values, and the model with the lowest average of these three criterion was chosen as the best overall model using this method. This model also happened to be the same model chosen to be optimal in the analytical approach, being SARIMA$(1,1,1)$ $(1,0,1)_4$. Therefore, the data-driven approach validated the analytical approach and strengthened confidence in the chosen model as the optimal model for the series.

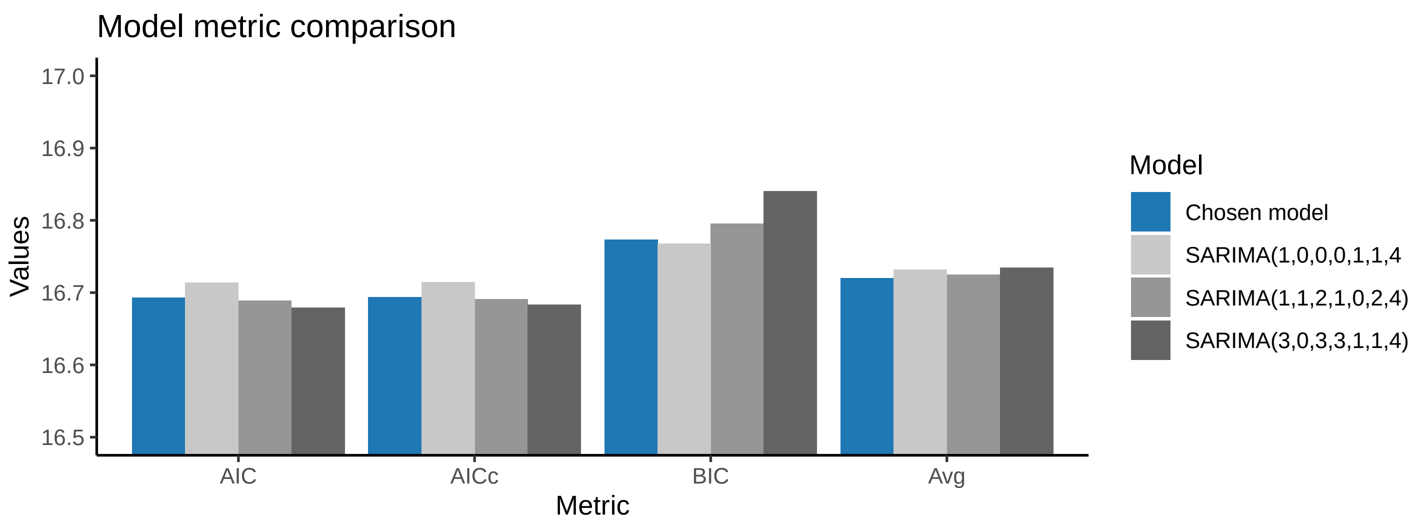
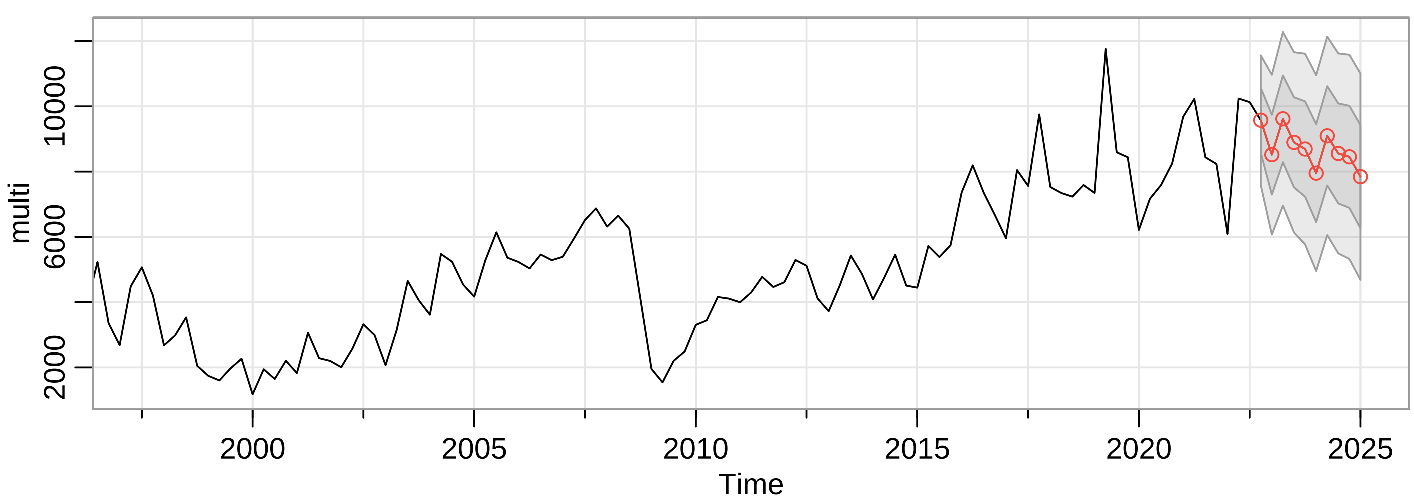


Figure 7: Model metric comparison

Predictions



The plot shows the next 10 predicted observations for the housing starts data in British Columbia. The predicted values appear to be reasonable and realistic, and they show a slight downward trend. This suggests that the chosen model is accurately capturing the dynamics of the housing market in the province, and that short term predictions are likely to be reliable.

Conclusion

Advantages and drawbacks

The model appears to be effective at making short-term predictions, as indicated by the good fit of the model to the data. In addition, there is no evidence of overfitting, as indicated by the significance of the model parameters (i.e., the low p-values). These features make the model a useful tool for forecasting multi-family housing starts in the short-term.

However, the model has some limitations. It is not able to make accurate long-term predictions, as it does not account for long-term seasonality or other underlying trends in the data. Furthermore, the model does not take into account broader economic factors that may affect the housing market. As a result, the model may not be reliable for forecasting housing starts over the long-term or in situations where the economic environment is changing rapidly. In summary, while the model is good at making short-term predictions, it has some limitations that should be considered when using it to make decisions.

Concluding thoughts

In conclusion, economic factors appear to play a significant role in the construction of multi-family housing. These types of units require large capital investments[8], which may be more sensitive to economic downturns than other types of housing. As a result, the number of multi-family starts may be more volatile than the number of single-family starts. Furthermore, the data indicate that the construction of multi-family homes is less subject to seasonality than single-family homes, likely due to longer relative construction times, market demand, and other economic factors.

Overall, this analysis has provided valuable insights into the housing market in British Columbia. It has been shown that the popularity of multi-family housing has increased relative to other types of units, and that economic factors are important drivers of the construction of these units. By using both an analytical and data-driven approach, an effective model for predicting short-term housing starts was selected. In doing so, the objectives were fulfilled and useful information was provided for anyone interested in the housing market in British Columbia.

References

[1] Vancouver Sun - B.C. posts Canada's highest unaffordable housing rate, homeless fear death on street

[2] Washington Post - If policymakers are serious about tackling inflation, they need to address soaring housing costs

[3] CBC - McMaster grad students 'don't have enough money to buy food,' want university to increase funding

[4] Science Direct - The implications of condominium neighbourhoods for long-term urban revitalisation

[5] Bloomberg - The Relationship Between Skyscrapers and Great Cities

[6] Pendragon Homes - Why Spring Is A Great Time To Start Building A Home

[7] Remax - The Difference Between Condos and Apartments

[8] ProEst - How Much Does It Cost To Build An Apartment Complex?

Appendix - R Code

Setup

```r
# Setup
knitr::opts_chunk$set(echo = F, results = "hide", warning = FALSE)
library(astsa)
library(ggplot2)
library(knitr)

MSE <- function(resids) {
    return(mean(resids^2))
}
```