Forecast Reconciliation for Land Transfer Duty

Hoang Do

Report for
The Department of Treasury and Finance

1 August 2024

1 Abstract

Accurate land transfer duty forecasting is essential for effective government market monitoring and policy implementation. This report presents a novel forecasting methodology for land transfer duty (LTD) in Victoria, emphasizing the application of forecast reconciliation techniques. Time series data can often be disaggregated by various attributes of interest. And forecasts are frequently needed for both disaggregated and aggregated series, necessitating that the forecasts align accurately with the data's aggregation structure. The forecast reconciliation methodology addresses this challenge by ensuring coherence across the entire aggregation structure. By employing Vector Error Correction Model (VECM) within cross-sectional, temporal, and cross-temporal hierarchical structures, we find that the methodology significantly improves forecast accuracy of Victoria's land transfer duty, using the *Department of Treasury and Finance*'s forecasts as a benchmark. Moreover, the methodology is broadly applicable when the aggregate variable of interest can be modelled for disaggregated components and/or at different frequencies.

2 Introduction and background

The property sector plays a pivotal role in Australia's economy, accounting for 1 in 4 jobs indirectly and contributing around 13% of Gross Domestic Product (GDP). In the 2021 financial year, property sales totaled approximately $350 billion as stated in **reia_2022**. Land transfer duty, previously known as stamp duty, significantly impacts property transactions and the sector as a whole. A study published by the New South Wales Treasury found that a 100 basis point (1%) cut in land transfer duty could boost property transactions by 10% (**nsw_treasury_2021**).

Land transfer duty is a tax applied to the "dutiable value" of a property being purchased or acquired, whether it is a first home or an investment property. The dutiable value is determined as either the property's purchase price or its market value, whichever is greater. Several factors influence the amount of duty paid, including the buyer's intended use of the property, foreign purchaser status, and eligibility for exemptions.

Additionally, Victoria's tax revenue heavily relies on land transfer duty. As shown in Figure **??**, over the past 20 years, on average, land transfer duty accounts for **27%** of Victoria's tax revenue. Despite its perceived inequity and various exemptions designed to aid homebuyers, abolishing this duty remains challenging due to the need for equivalent revenue replacement. If this duty is removed, the government will need to introduce one or more new taxes to generate equivalent revenue.

To balance the need for sufficient tax income without discouraging property transactions, *the Department of Treasury and Finance* (**DTF**) has been supporting the provision of better advice to the
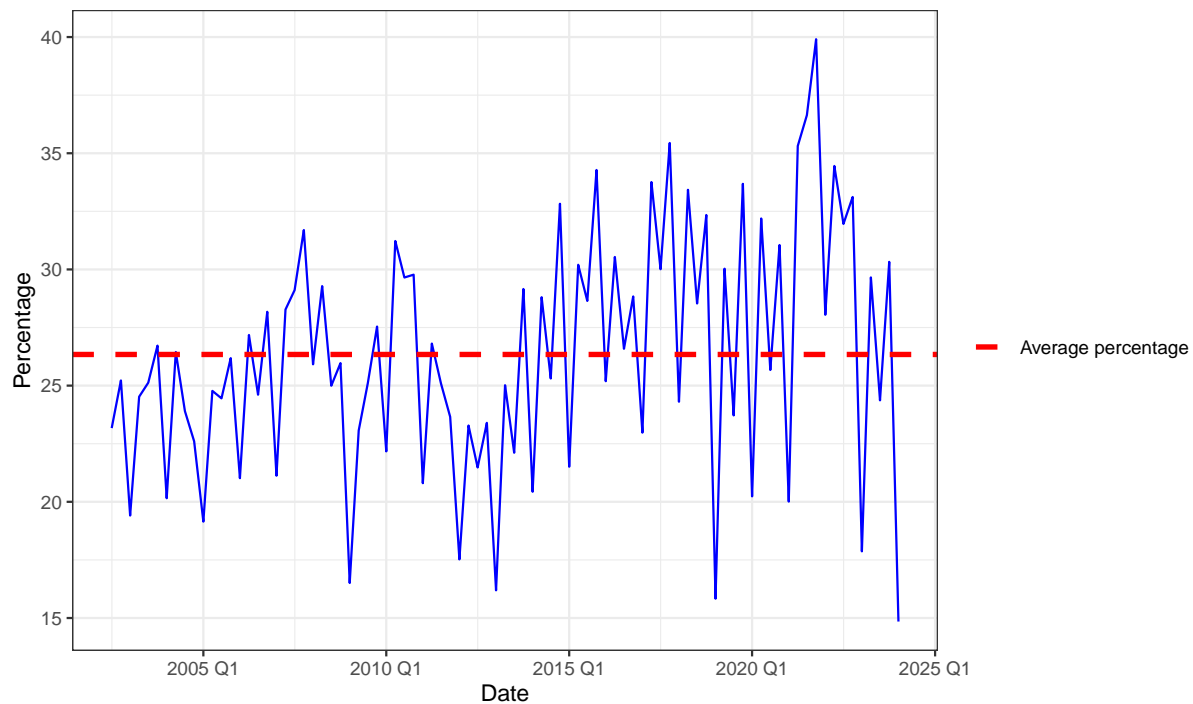
Figure 1: *Percentage of Victoria's Tax Revenue from Land Transfer Duty*

government on policy adjustments, subsidies, exemptions, and restrictions by delivering evidence-based insights. Accurate forecasts of land transfer duty, both short-term (*1 to 3 months*) and long-term (*12 months*), are essential for these decisions.

This report introduces a new forecasting methodology aimed at improving the accuracy of these predictions. By employing forecast reconciliation and combining cross-sectional and temporal hierarchies, we aim to provide the *Department of Treasury and Finance* with more reliable forecasts to inform policy-making. The primary advantage of this methodology is that it produces forecasts at various aggregation levels within both cross-sectional and temporal hierarchies, allowing for the inclusion of unique characteristics of different nodes. Our results use monthly land transfer duty data, covering the period from July 2013 to March 2024.

Our project employs three forecasting models, including vector error correction model (VECM), vector autoregressive model (VAR) and autoregressive integrated moving average (ARIMA). For the model fitting process, we include three macroeconomic indicators and property market indices, which are sales, home value index and lending provided by the Australian Bureau of Statistics (ABS) and CoreLogic. The rationale for selecting these three explanatory variables is that the best-performing models used by the *Department of Treasury and Finance* also employ them.

To assess the forecast performance, we use time series cross-validation to ensure that the corresponding training set consists only of observations that occurred prior to the observation that forms the test set.

The length of the first training set contains 108 months of data (equivalent to 9 years) and increase the size of successive training sets by one month. As a result, there will be 10 testing sets with 1 to 12-step-ahead forecasts. We then computed accuracy measures, specifically root mean squared errors (RMSE) and mean absolute percentage errors (MAPE), for all forecasts, and finally averaged them across the 10 sets.

Overall, the results suggest that VECM, particularly when using cross-temporal reconciliation, consistently outperforms ARIMA. Temporal hierarchies contribute more to accuracy improvements than cross-sectional hierarchies. The reconciled forecasts from VECM exhibit lower RMSE and MAPE compared to the *Department of Treasury and Finance*'s forecasts, underscoring the efficacy of the proposed approach.

The rest of the paper is organized as follows. Section 3 includes exploratory data analysis (EDA). Section 4 explores the forecasting methods covered in detail. Section 5 discusses the detailed time series cross-validation process adopted for this project. Section 6 presents the main results and their implementation. Section 7 concludes the study and offers recommendations.

3   Exploratory Data Analysis (EDA)

In this section, we conduct a detailed Exploratory Data Analysis (EDA) for the total land transfer duty and all its disaggregated levels, as well as for other included exploratory variables. Each of the three exploratory variables serves as a measure of property market performance and will have interrelations with land transfer duty.

3.1   Time series analysis

For this project, the total land transfer duty is divided into residential and non-residential categories. The non-residential land transfer duty is further disaggregated into commercial, industrial, and other categories. Figure **??** provides an overview of the cross-sectional hierarchical structure.

Total Land Transfer Duty

Figure **??** demonstrates an upward, non-linear trend, showing that the value of land transfer duty has generally increased over the ten-year period. Note that, due to the outbreak of Covid-19, total land transfer duty experienced a significant decline, reaching a low of 286.2 million in June 2020. Subsequently, it rebounded to an unprecedented peak of 1.013 billion in March 2022. This trend can be attributed to multiple rate cuts by the Reserve Bank of Australia, which initially aimed to stabilize the real estate sector and subsequently stimulated demand. Additionally, there is a noticeable increase in variability, particularly in recent years, indicating more pronounced fluctuations in the values. This
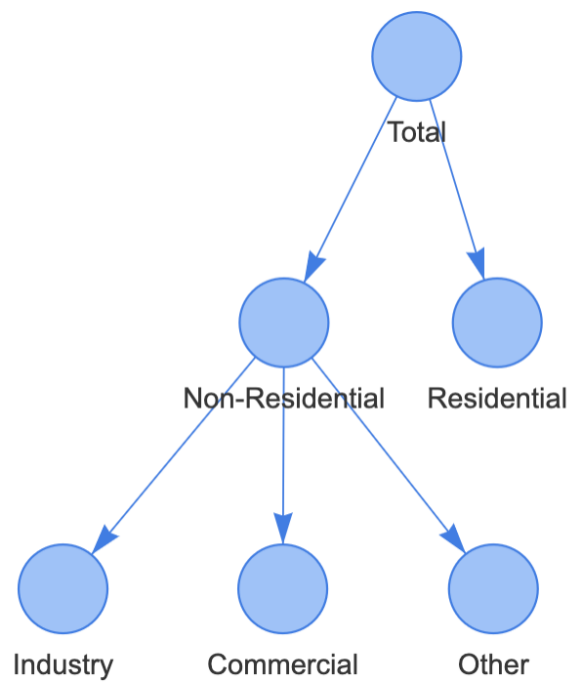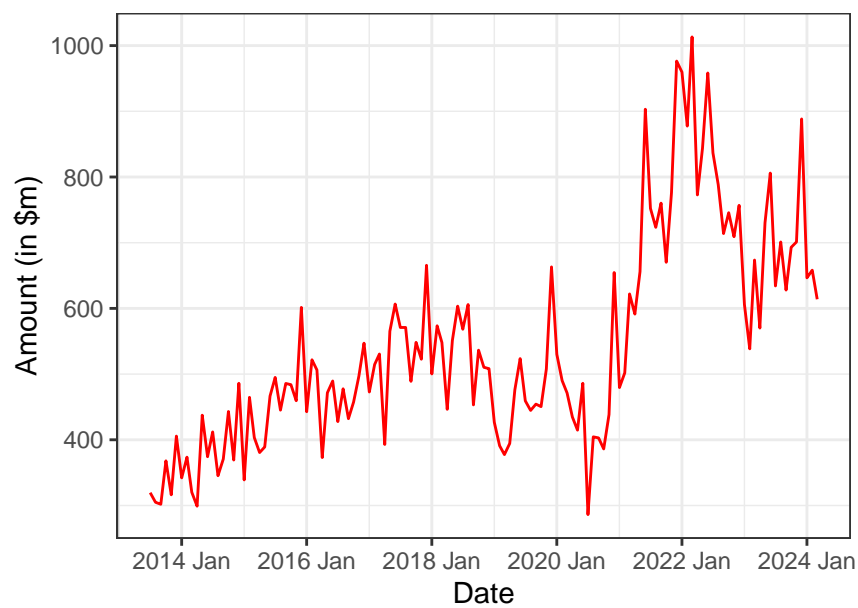
Figure 2: *Cross-sectional hierarchy*

Figure 3: *Time plot of Land Transfer Duty in Victoria over time*

increasing volatility suggests that a transformation of the data may be necessary to better analyze and model these trends.
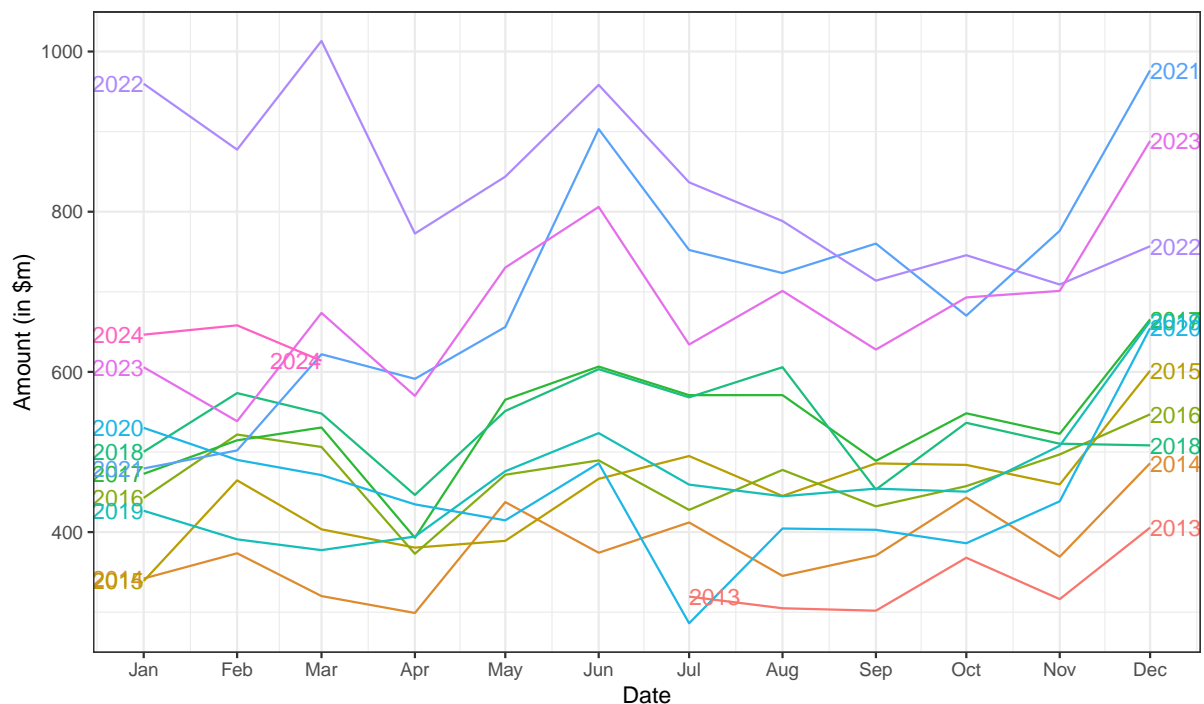

Figure 4: *Seasonal plot of Land Transfer Duty in Victoria*

Considering the seasonal patterns of land transfer duty, referring to Figure **??**, it is clear that there is a large jump in land transfer duty in June and December each year. This may be attributed to June being the end of financial year, when buyers rush to complete transactions to take advantage of tax benefits, financial reporting, while December being the end of calender year, when many individuals aim to complete transactions before the holiday season. Figure **??** also shows that there is a decrease in April each year.

Figure **??** shows that there is a seasonal pattern for log transformed land transfer duty, which agrees with what indicated from Figure **??**. The grey bars positioned to the left of each panel illustrate the relative scales of the components. Although each grey bar represents an same length, their sizes differ due to the distinct scales of the plots. The prominent grey bar in the seasonal panel indicates that the variation in the seasonal component is the smallest when compared to the overall variation in the data.

Moreover, Figure **??** demonstrates that after applying a log transformation, the variability becomes more consistent. However, the relatively small grey bar to the left of the remainder panel indicates that some heterogeneity persists.

The grey line in Figure **??** represents the log-transformed total land transfer duty, while the blue line
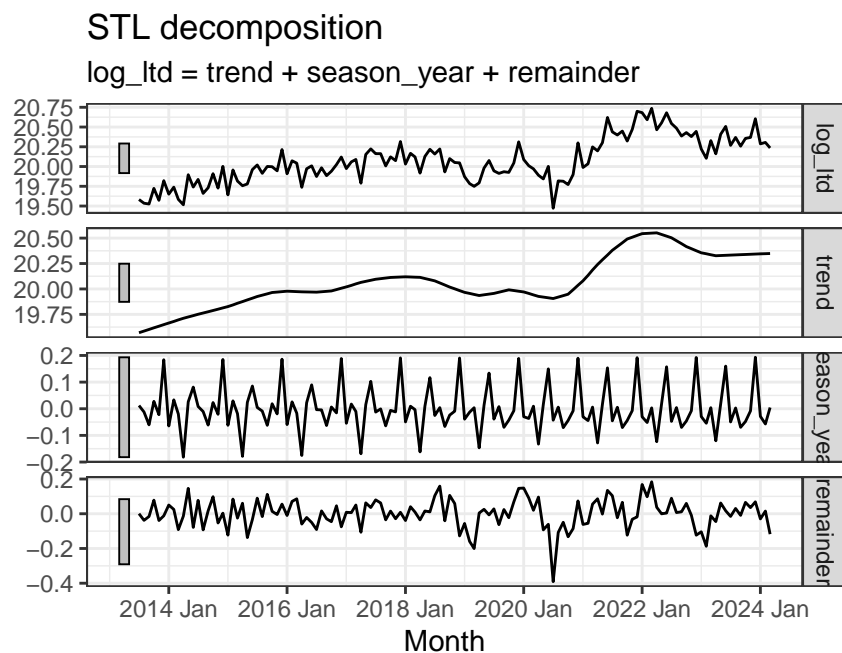
Figure 5: *Log of total Land Transfer Duty in Victoria (top) and its three additive components.*
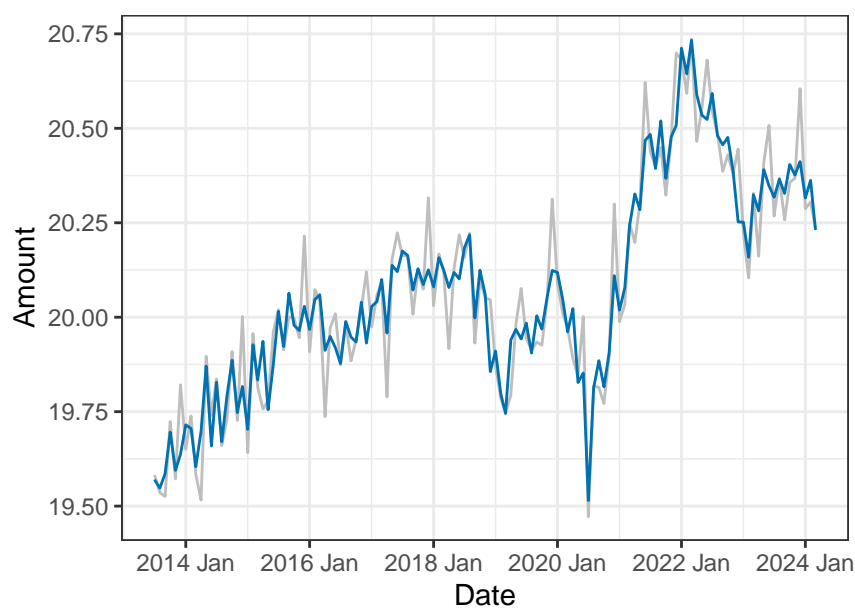

Figure 6: *Seasonally adjusted of log Land Transfer Duty (blue) and the original data (grey)*

represents the seasonally adjusted log-transformed total land transfer duty. Figure **??** suggests that by adjusting for the seasonal pattern, some local peaks and troughs have been smoothed out.

Residential property land transfer duty

Figure **??** illustrates a non-linear upward trend in residential property land transfer duty, closely mirroring the overall trend in total land transfer duty. The figure also reflects the impact of the
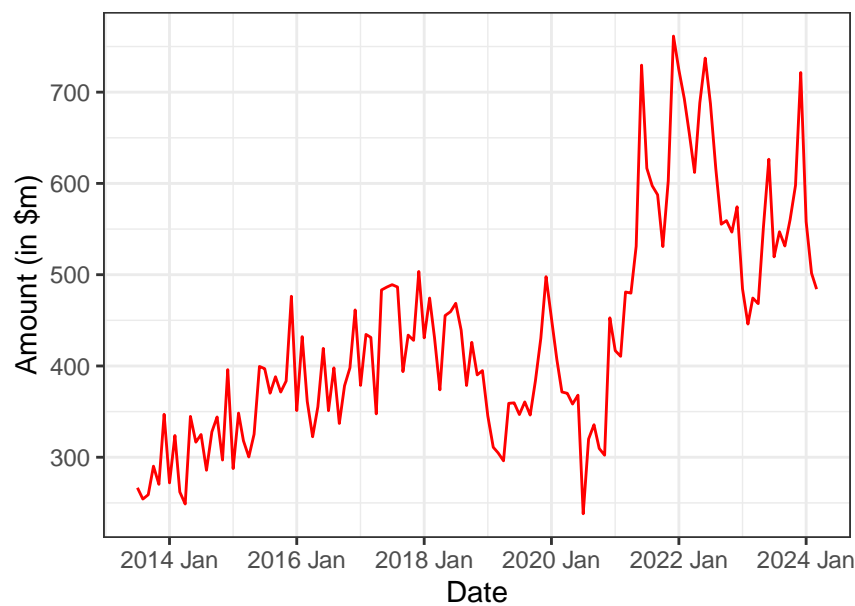
Figure 7: *Time plot of Residential property Land Transfer Duty in Victoria*

Covid-19 pandemic, with a significant decline to a low of 238.1 million in June 2020, followed by a rebound to an all-time high of 761.4 million in December 2021. This pattern is attributable to the fact that residential property comprises the largest portion of land transfer duty. These findings suggest that residential property remains a crucial component of the market, likely reflecting trends in housing demand and price fluctuations. Additionally, there is an increase in variability over time, which also suggests a transformation of the data.

Since as discussed above, we expect residential land transfer duty also mirrors the seasonal pattern of total land transfer duty, which is shown in Figure **??**. There is also a large jump in land transfer duty in June and December each year, which can be explained by the same reason.

Non-residential property land transfer duty

Figure **??** shows a markedly different pattern in non-residential property land transfer duty compared to total and residential property duty. While non-residential duty reached an all-time high of 357.2 million in March 2022 post-Covid-19, it lacks a consistent long-term growth trend, exhibiting significant variability. Moreover, the sudden drop after reaching its all time high in March 2022 may suggest that non-residential sector reacts more significantly to changes in the cash rate than the residential sector.

Figure **??** also shows an increase in June and December each year. Although the magnitude of change is smaller compared to total and residential property land transfer duty, the increasing pattern is more consistent across all years.
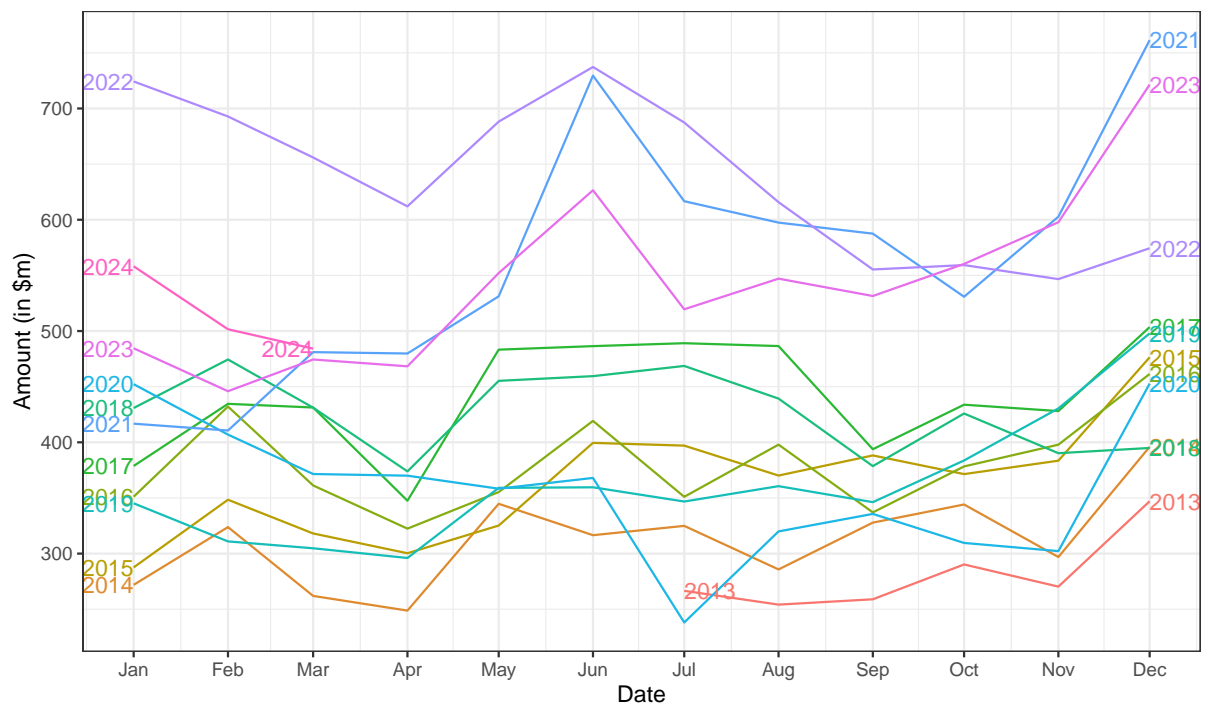
Figure 8: *Seasonal plot of Residential property Land Transfer Duty in Victoria*
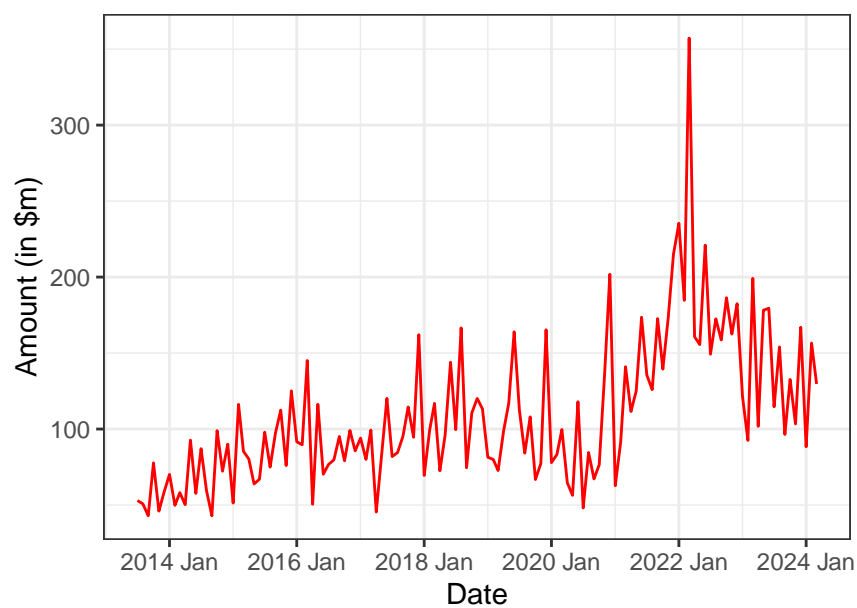

Figure 9: *Time plot of Non-residential property Land Transfer Duty in Victoria*
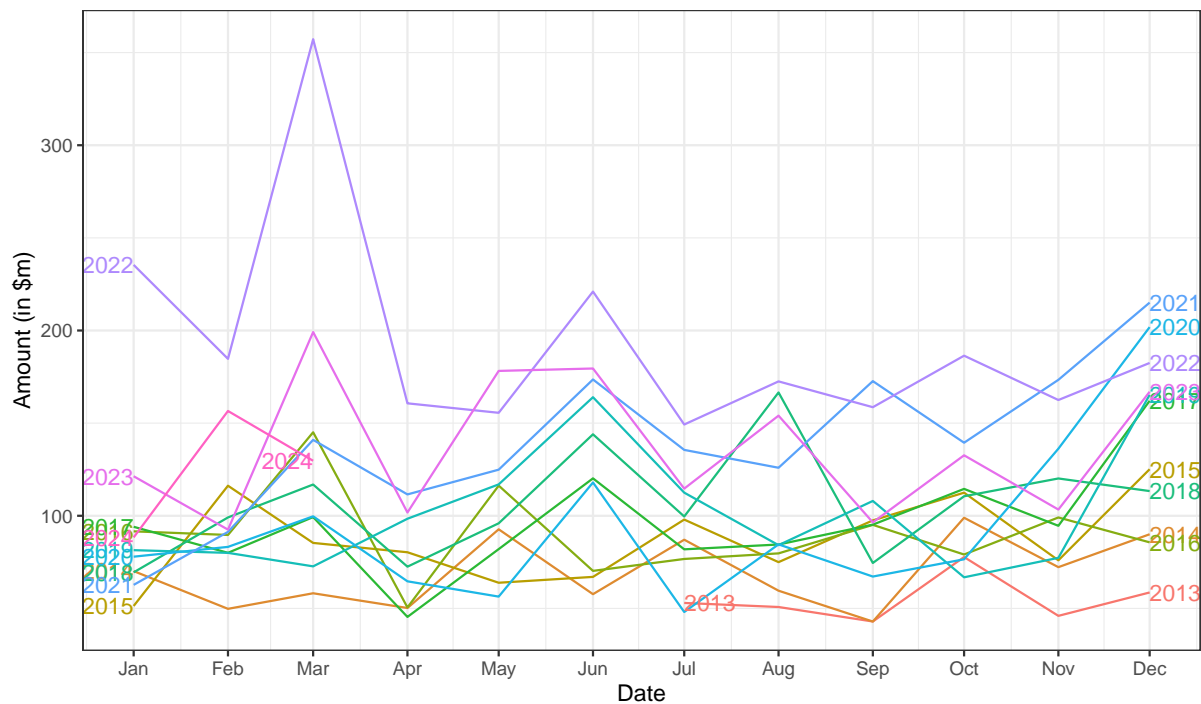
Figure 10: *Seasonal plot of Non-residential property Land Transfer Duty in Victoria*

Commercial property land transfer duty

Figure **??** demonstrates the fluctuations in commercial property land transfer duty from 2014 to 2024. The plot also exhibits significant variability, with several peaks and troughs over the observed period, ranging from 22.1 million to 132.4 million. Notably, the commercial land transfer duty reached its highest levels in early 2022, reflecting a post-Covid-19 recovery. However, unlike the residential sector, the commercial property sector does not show a consistent long-term growth trend.

Figure **??** demonstrates various increases throughout the year, in March, May, June and December, highlighting significant fluctuations throughout the year. Moreover, while some years display consistent seasonal patterns, others show irregular fluctuations.

Industrial property land transfer duty

Figure **??** shows a more stable variability compared to residential or total land transfer duty, with less pronounced fluctuations, apart from notable peaks observed in mid-2018 and March 2022. Overall, the trend is non-increasing, indicating a lack of long-term growth. The industrial sector has collected land transfer duty between 5.2 million and 214.4 million, highlighting some large-scale industrial transactions or developments during certain periods. These characteristics suggest that while industrial property transactions exhibit periodic spikes, they generally maintain a steady pattern without significant upward or downward trends.
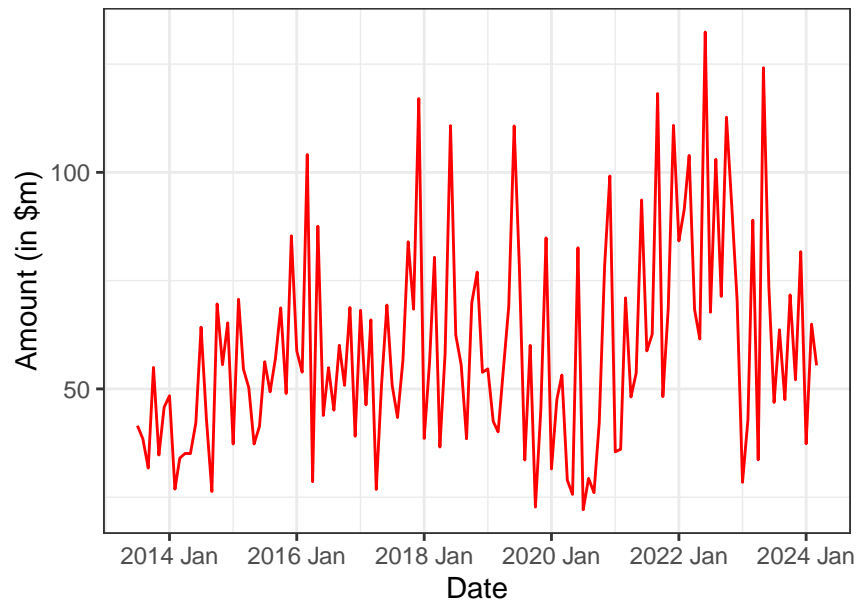
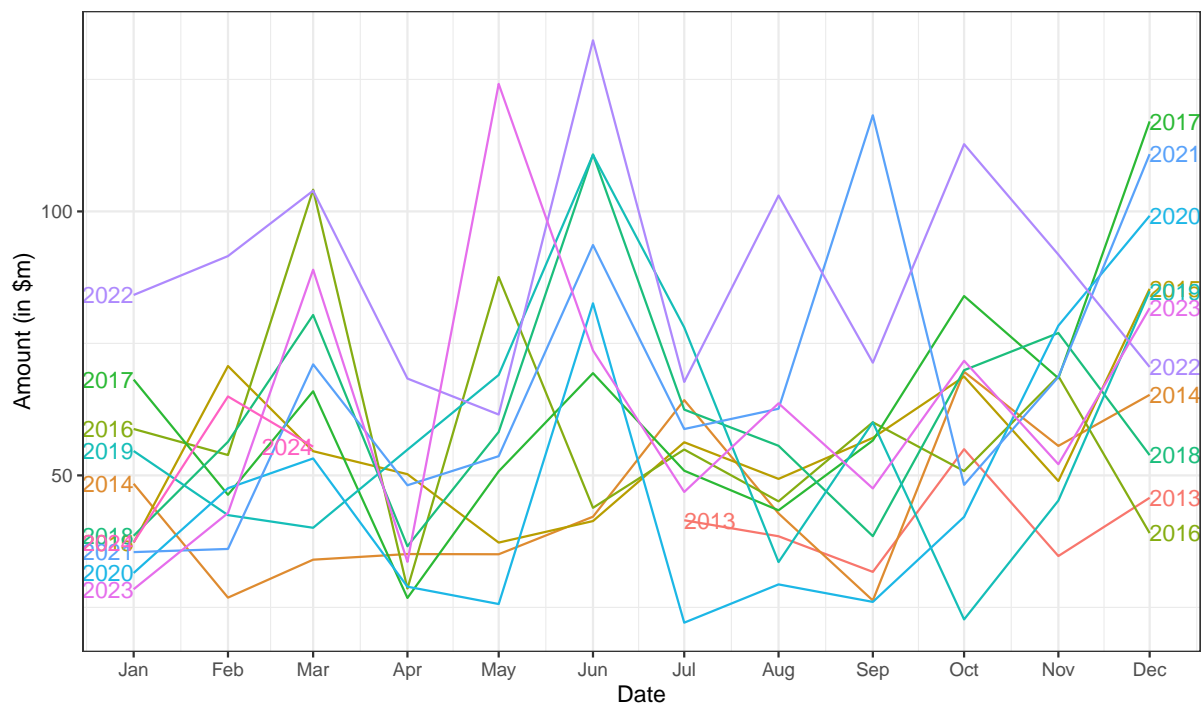**Figure 11:** *Time plot of Commercial property Land Transfer Duty in Victoria*

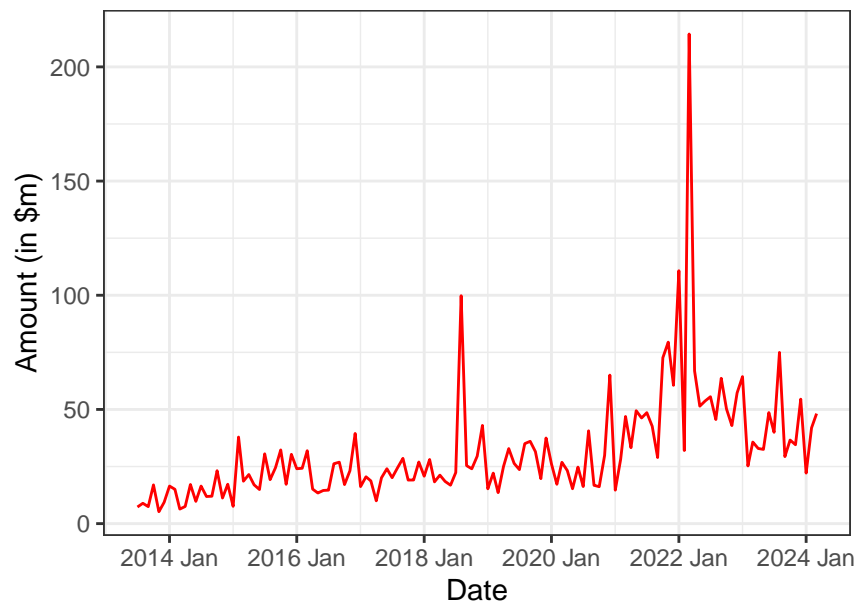**Figure 12:** *Seasonal plot of Commerical property Land Transfer Duty in Victoria*

Figure 13: *Time plot of Industrial property Land Transfer Duty in Victoria*
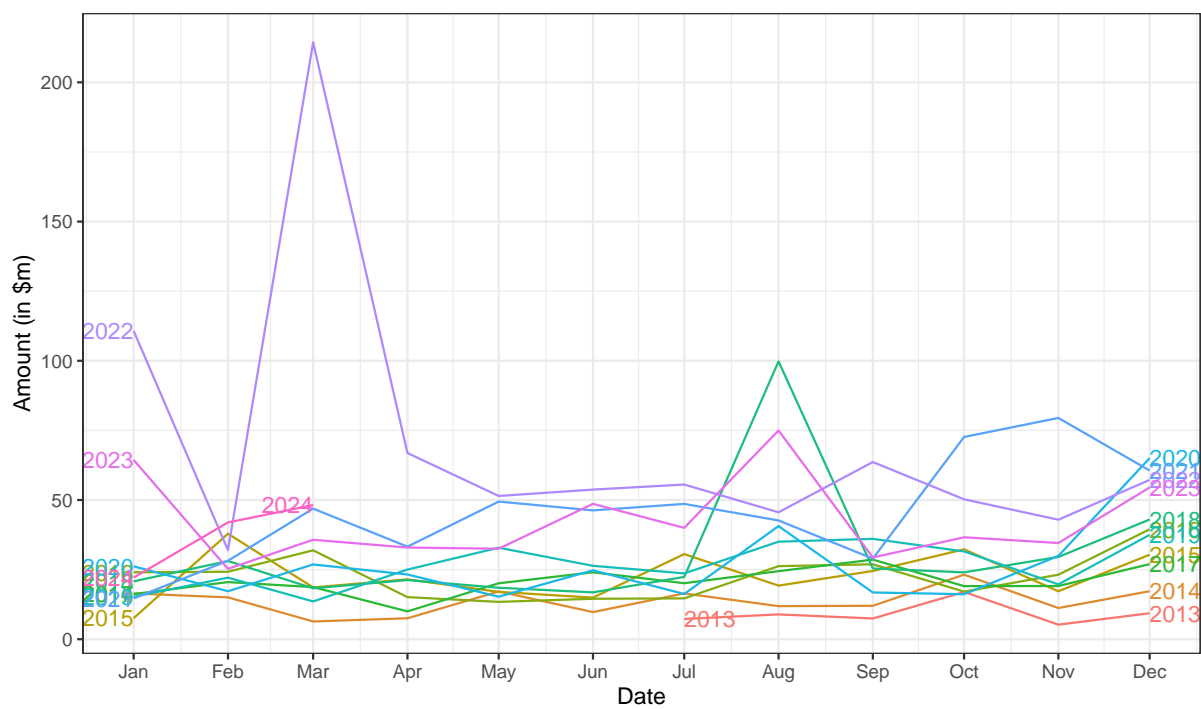

Figure 14: *Seasonal plot of Industrial property Land Transfer Duty in Victoria*

Apart from slight increases in August and December each year, Figure **??** shows that there is no clear indication of seasonal pattern.
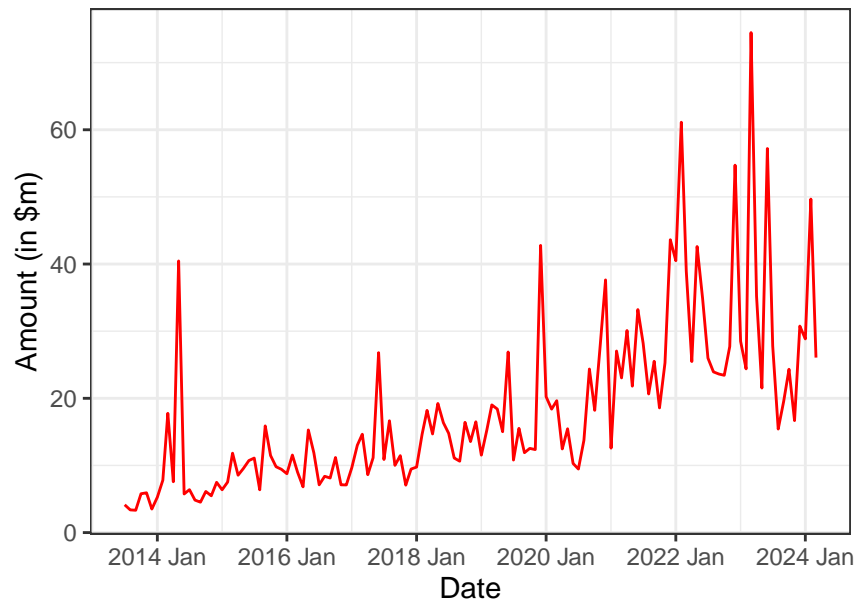
Other property land transfer duty

Figure 15: *Time plot of Other type property Land Transfer Duty in Victoria*

Figure **??** shows a non-linear increasing trend in other type property land transfer duty, including agricultural with different variability over time and irregular peaks. Contributions from these sectors range from 3.3 million to 74.5 million, indicating sporadic activity that may correspond with specific market or economic conditions.

As shown in Figure **??**, the seasonal pattern is quite irregular across all years, but there is a consistent increase observed in December for the past several years from 2020.

Sales

The sales variable indicates the number of properties sold in Victoria, measured in units at the contract date for all types of residential dwellings.

Figure **??** shows a non-linear trend, and strong fluctuations in sales units over the ten-year period from January 2014 to January 2024. The plot exhibits considerable variability with both high and low peaks scattered throughout the timeline. The values range from approximately 5000 to over 15000, indicating significant changes in sales units.

In recent years, there is a noticeable increase in variability, particularly around 2020 and onwards. This increasing volatility indicates that a transformation of the data may be necessary to better analyze
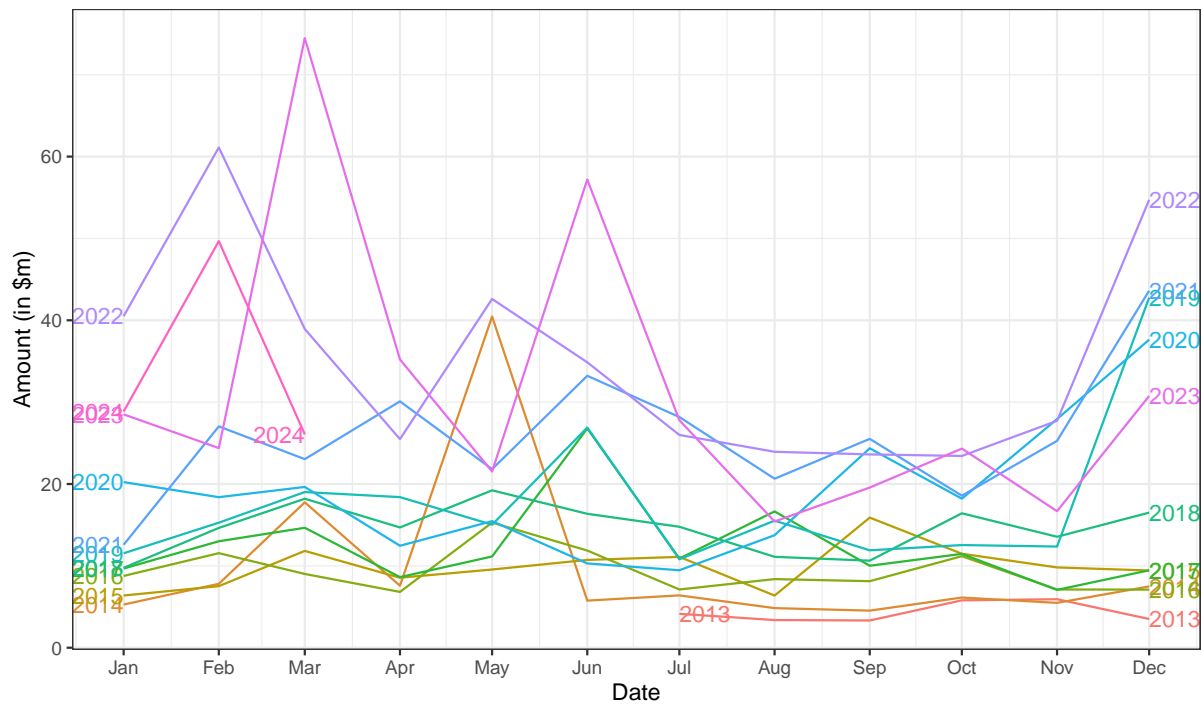
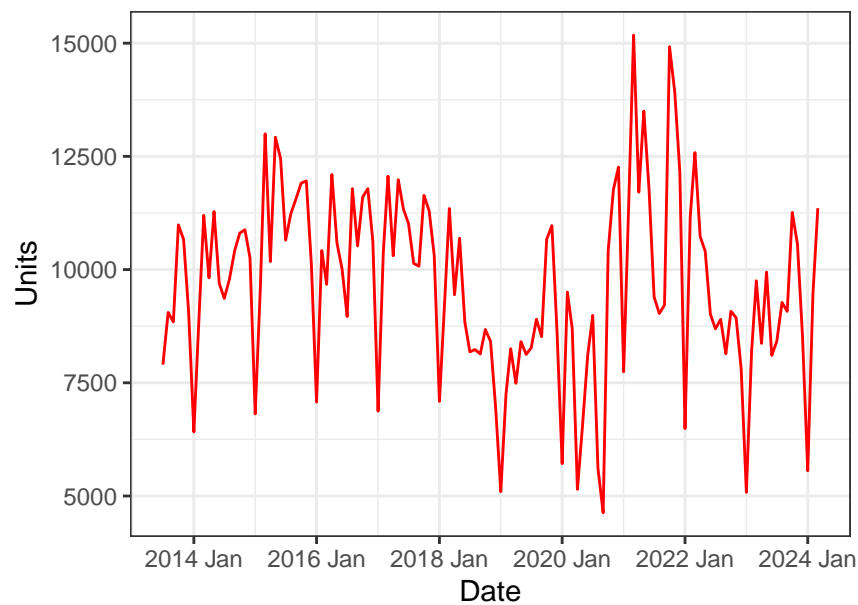**Figure 16:** *Seasonal plot of Other type property Land Transfer Duty in Victoria*

Figure 17: *Time Plot of Units of Properties Sold in Victoria*

and model these trends.
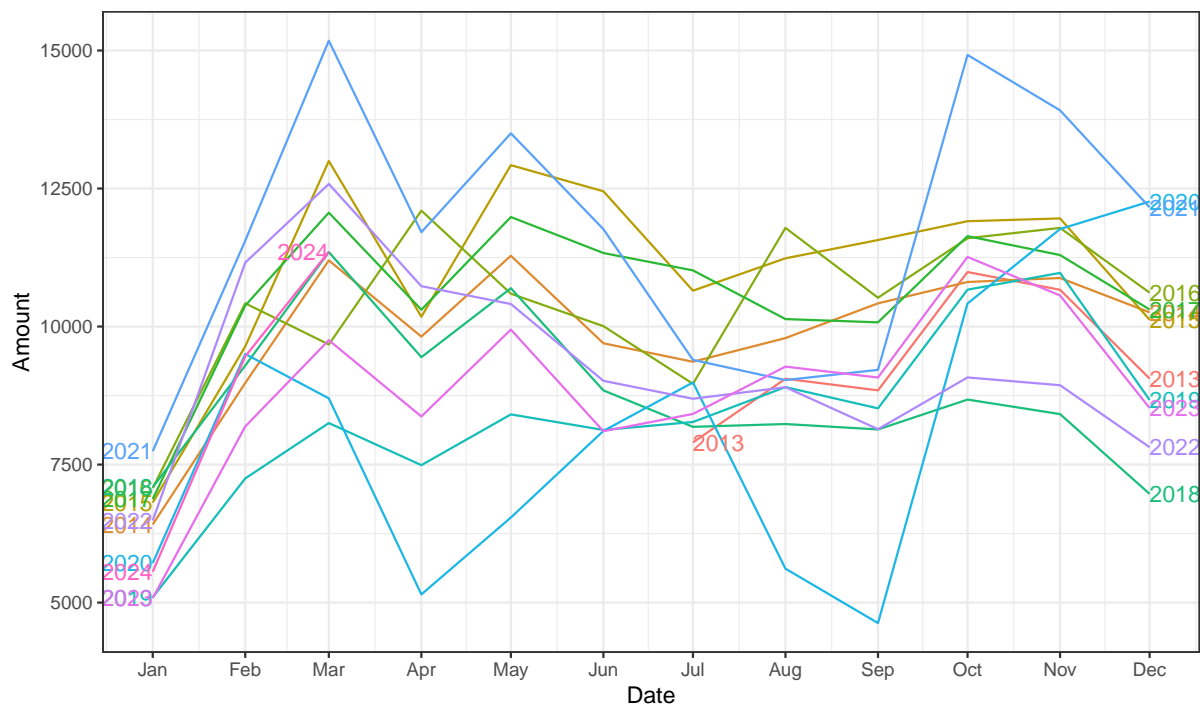
Figure 18: *Seasonal plot of Units of Properties Sold in Victoria*

Figure **??** exhibits a significant decline in sales units in December, followed by a trough in January each year. This trend contrasts with the observed increase in total land transfer duty in December, as depicted in the seasonal plot in Figure **??**. This discrepancy can be attributed to the fact that, despite the reduction in the number of sales units, the properties sold during this period are typically of higher value, thereby driving up the total land transfer duty.

Figure **??** indicates a seasonal pattern for log-transformed sales, which agrees with Figure **??**. However, the large grey bar in the seasonal panel indicates that the variation in the seasonal component is the smallest when compared to the overall variation in the data.

Moreover, Figure **??** demonstrates that after applying a log transformation, the variability becomes more consistent, although some heterogeneity still remains.

The grey line in Figure **??** represents the original log-transformed sales, while the blue line represents the seasonally adjusted log-transformed sales. Figure **??** suggests that there is a huge difference between seasonally adjusted sales and original sales, especially in periodic drops.

Home Value Index

The Home Value Index aims to measure monthly movements in the value of Australian housing markets.
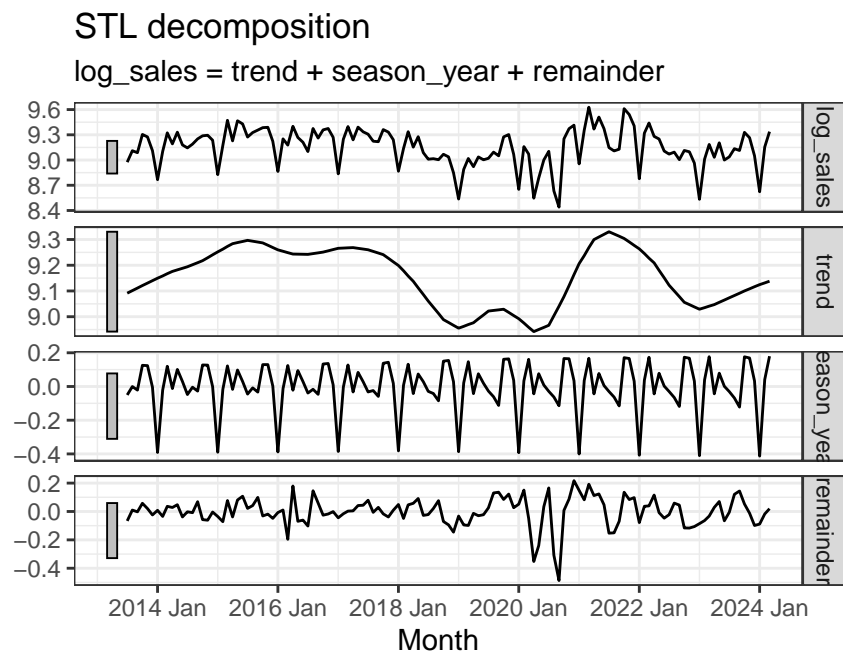
Figure 19: *Log of Units of Properties Sold in Victoria (Top) and Its Three Additive Components*
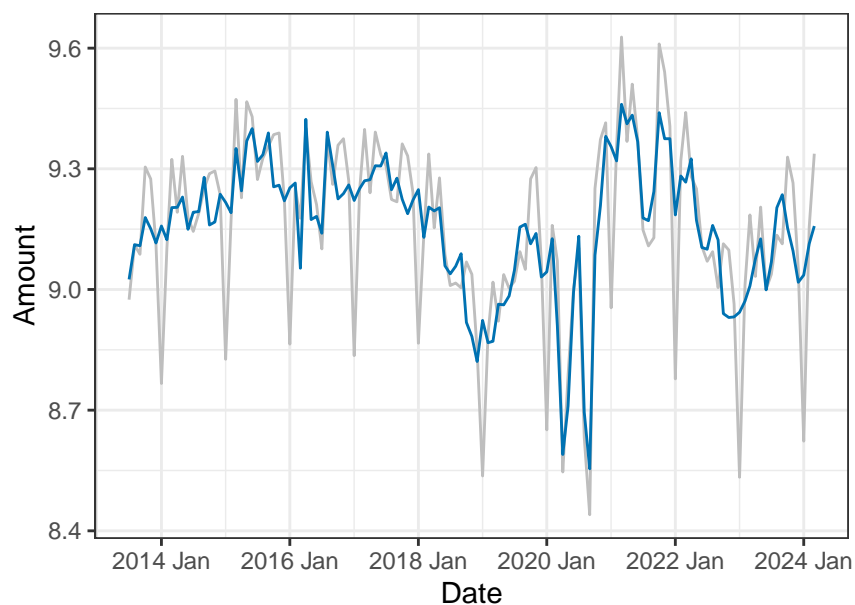
Figure 20: *Seasonally adjusted of log of Units of Properties Sold in Victoria (blue) and the original data (grey).*
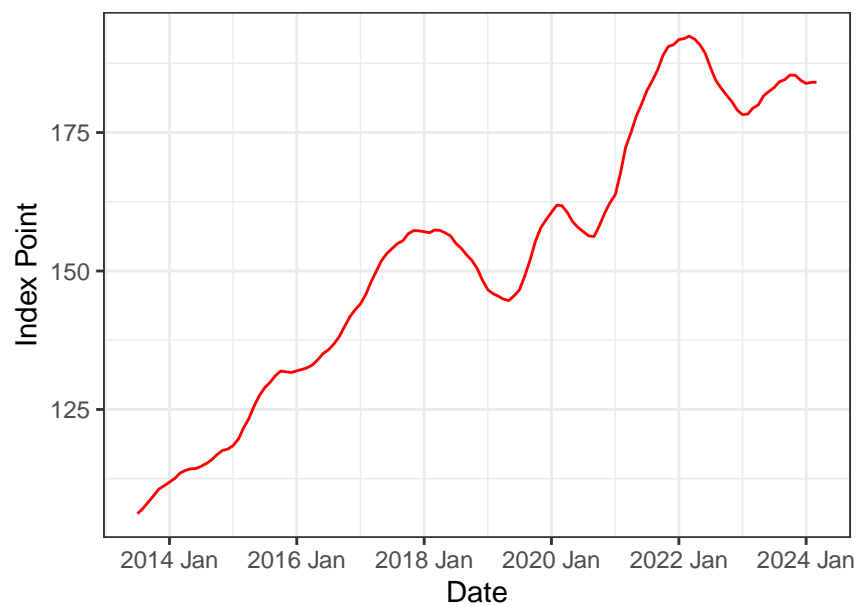
Figure 21: *Time plot of Home Value Index in Victoria over time*

Figure **??** illustrates an increasing trend, with some considerable variability throughout the timeline. There is no clear sign of seasonal pattern in home value index.

Lending

Lending is the new borrower-accepted finance commitments for housing, personal and business loans from Australian Bureau Statistics (ABS), relating to all types of residential transactions.
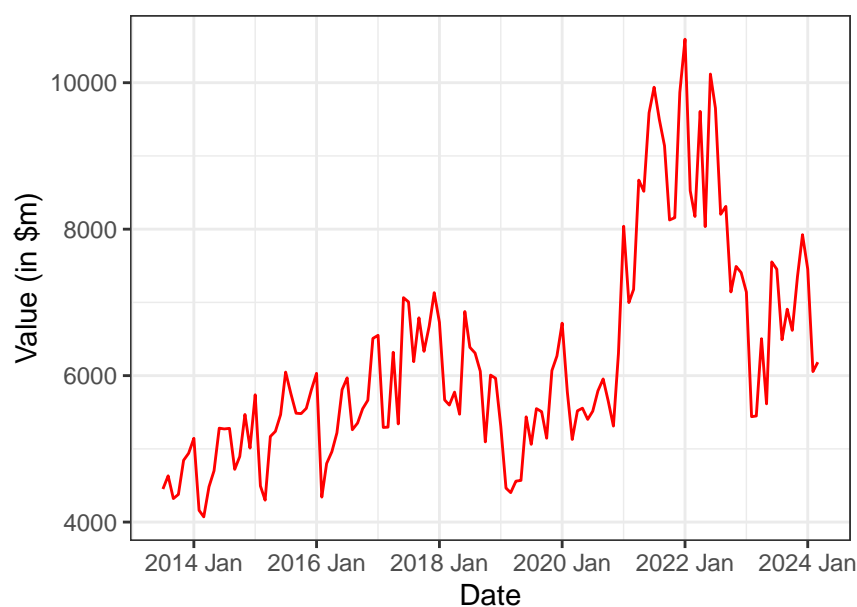
Figure 22: *Time plot of Lending in Victoria over time*

Table 1: *Johansen test result*

	Rank	Test Statistic	10%	5%	1%
r <= 2 |	r <= 2	2.30	6.50	8.18	11.65
r <= 1 |	r <= 1	17.34	15.66	17.95	23.52
r = 0 |	r = 0	57.15	28.71	31.52	37.22

Figure **??** shows an overall increasing trend over time with clear seasonal pattern for lending.

3.2 Dealing with seasonality

As discussed in Section **??**, it is evident that many time series exhibit seasonal patterns. Given the limitations of Vector Error Correction Models (VECM) in adequately modeling these seasonal fluctuations, we propose a methodological adjustment. Specifically, we will decompose the time series into seasonally adjusted components and explicit seasonal factors. This decomposition allows us to apply VECM to the seasonally adjusted data. Subsequently, we will conduct independent forecasts for both the seasonally adjusted data and the seasonal components. The final step involves the reconciliation of these forecasts then integrating them.

3.3 Cointegration Analysis for modelling

Based on the observations from Figure **??**, which illustrates the trend and increasing variance pattern in land transfer duty while other disaggregated levels of land transfer duty exhibit either a trend or seasonal pattern, we can conclude that these series are non-stationary. The same findings apply to the three exploratory variables.

Furthermore, to justify the use of the Vector Error Correction Model (VECM) by the *Department of Treasury and Finance*, it is essential to test for the presence of cointegration patterns (or long-term equilibrium relationships) among these three series.

The Johansen procedure with the trace test statistic will be employed to test for cointegration between log transformed of land transfer duty, sales, and the home value index series. This procedure tests the null hypothesis that no cointegration relationship exists among the series.

The test results, as shown in Table **??**, using 3 lags as chosen by the DTF, allow us to reject the null hypothesis that $r \leq 1$, yet we fail to reject the null hypothesis that $r \leq 2$ at 5% level of significance. This suggests that among the three variables, the rank of the matrix exceeds 2, indicating the presence of at least two cointegration relationships.

Furthermore, a rank of 2 implies that we need a combination of at least two time series to form a stationary series.

To create such a linear combination, we can utilize the components of the eigenvector associated with the largest eigenvalue. According to the Johansen test summary, the largest eigenvalue is approximately 0.27. This corresponds to the eigenvectors components of `ltd_total`, which is approximately equal to (1, -1.16, -1.39). By forming a linear combination of the series using these eigenvector components, we can achieve a stationary series.
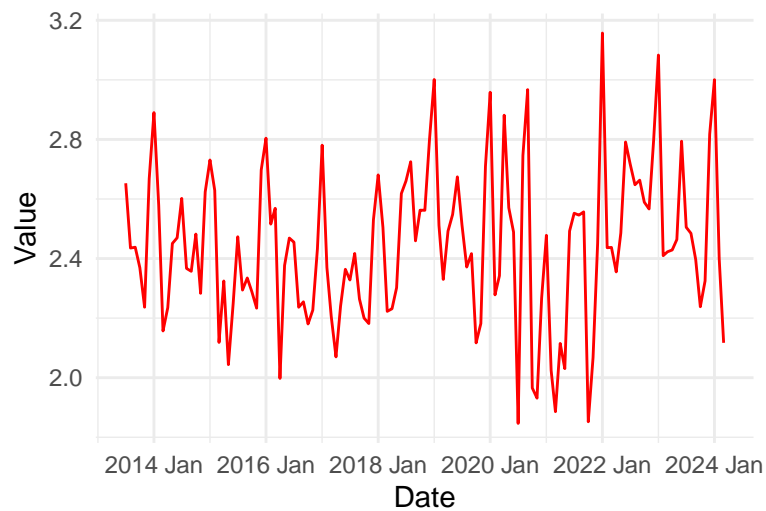

Figure 23: *Stationary series formed via a linear combination of 3 time series*

As shown in Figure **??**, the linear combination appears to be more stationary, although there remains a slight indication of varying variances over time.
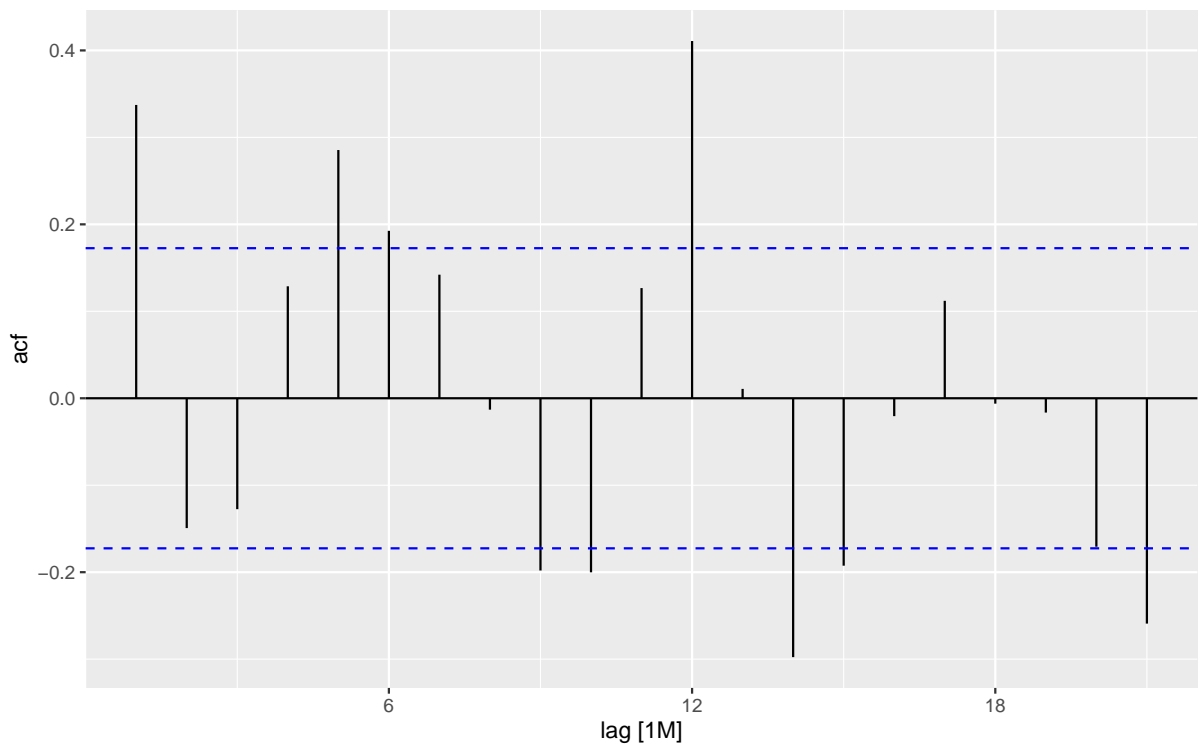
Table 2: *Johansen and ADF test for other type ltd*

Variable	Johansen p-value	ADF p-value
ltd_res	2	0.06
ltd_nonres	2	0.23
ltd_comm	2	0.14
ltd_ind	2	0.17
ltd_other	2	0.13

Additionally, the p-value of 0.09 from the Augmented Dickey-Fuller (ADF) test indicates that we can reject the null hypothesis of a unit root, providing evidence of a stationary series formed from the linear combination at the 10% level of significance. Therefore, we can conclude that it is appropriate to fit a Vector Error Correction Model (VECM) to these three time series.

Since we will fit VECM to all the variables in the hierarchy, including residential, non-residential, commercial, property and other, we will perform the same procedure of cointegration analysis as above.

As shown in Table **??**, the null hypothesis that the linear combination is non-stationary cannot be rejected for the variables `ltd_nonres`, `ltd_comm`, `ltd_ind`, and `ltd_other`. This finding suggests that alternative variables may be required to appropriately fit the Vector Error Correction Model (VECM) for these specific property categories in land transfer duty.

4   Methodology

4.1   Cross-sectional, temporal and cross-temporal hierarchies

Upon analyzing the data characteristics and the disaggregation of LTD, it becomes evident that a three-level hierarchical structure can be established. There is utility in generating forecasts at various levels of aggregation, driven by diverse reasons and objectives. For example, each level of aggregation may exhibit distinct characteristics; for instance, transactions involving residential properties might vary from those involving non-residential properties due to differences in market dynamics or market size for each property type.

In this context, total land transfer duty is divided into two categories: residential and non-residential properties. Non-residential properties are further disaggregated into three sub-categories: commercial, industrial, and other, which predominantly includes agricultural properties. In addition to the total land transfer duty for all property types, the government and the *Department of Treasury and Finance* may also be interested in forecasts for each category and sub-category. Figure **??** illustrates the cross-sectional hierarchical structure of land transfer duty as discussed.
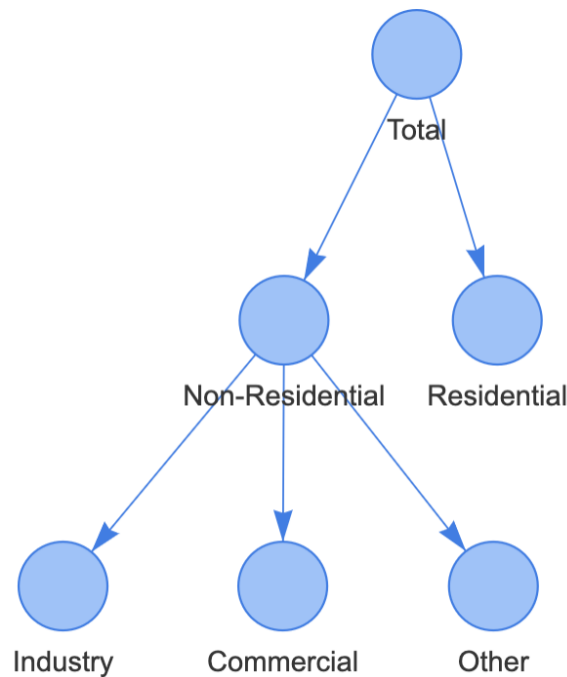
Figure 24: *Cross-sectional hierarchy*

Given that this is a straightforward cross-sectional hierarchical structure, we will also consider temporal hierarchies to further enhance forecast accuracy. Land transfer duty is collected monthly, and forecasts can be generated at bi-monthly, quarterly, four-monthly, semi-annual, and annual frequencies. Various temporal hierarchies can be constructed with monthly land transfer duty treated as the bottom level. Figure **??** presents an example of temporal hierarchies with monthly land transfer duty as the bottom level.
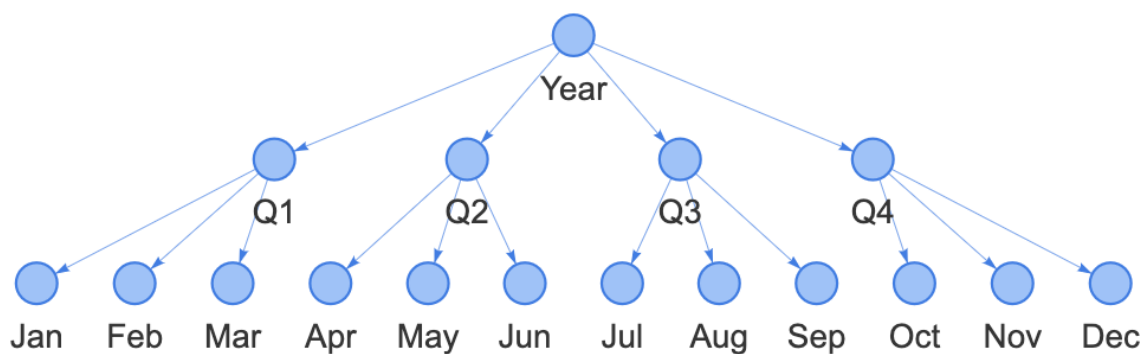
Figure 25: *Temporal hierarchy*

Although forecasts using cross-sectional and temporal hierarchies have demonstrated substantial improvements (**kourentzes2019cross**), these approaches have typically been used separately. By combining cross-sectional hierarchies and temporal hierarchies, referred to as cross-temporal hierar-

chies, forecast accuracy can be further improved. Moveover, another advantage of cross-temporal reconciled forecasts is that it provides aligned short-term and long-term decision. If cross-sectionally or temporally coherent forecasts are used disjointedly, some outputs may not be directly useful, such as very long-term forecasts at very disaggregate level, which is 2 years of forecasts at other type of property land transfer duty level. Therefore, one would have to post-process the forecasts further, for example combining together multiple long-term disaggregate bottom level forecasts (commerical, industrial, other) to produce long-term total land transfer duty forecasts, which would then break the desired coherence across all levels and time periods. Figure **??** from **kourentzes2019cross** shows an example of a cross-temporal hierarchical structure.
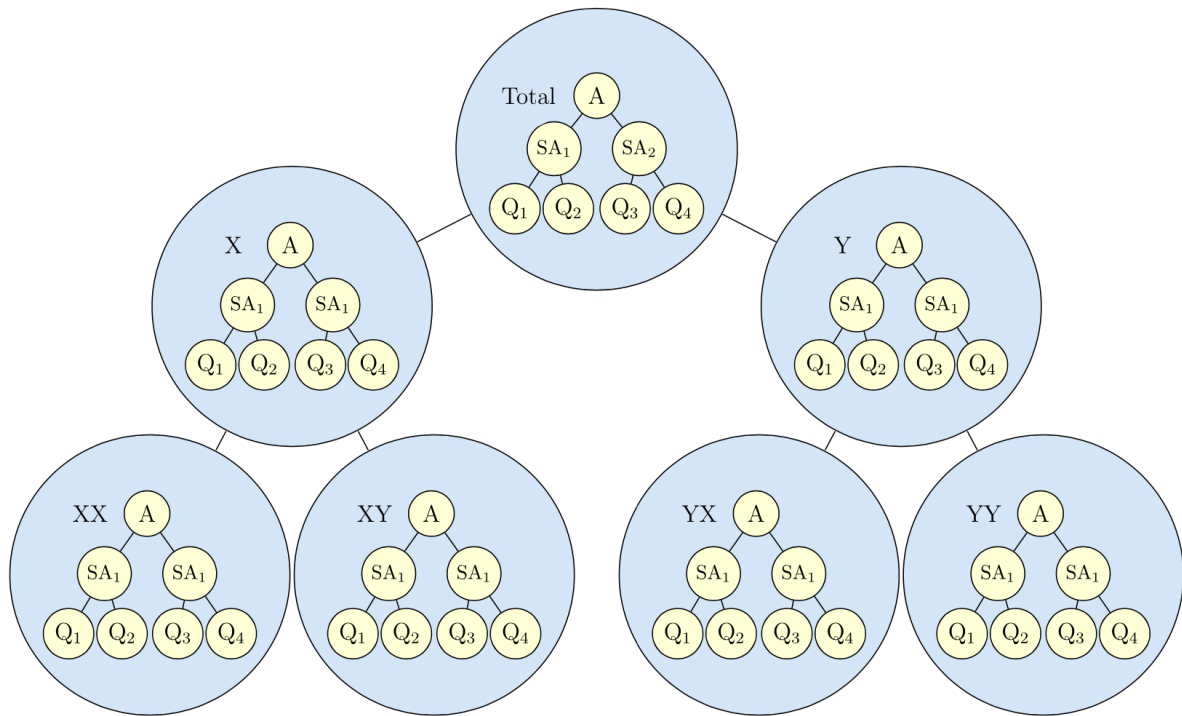
Figure 26: *Cross-temporal hierarchy*

4.2   Forecast reconciliation

From the cross-temporal hierarchies, forecasts can be produced at all levels for every nodes. And in an ideal scenario, forecasts from different levels of aggregation could seamlessly sum up to the top level. However, practical implementation often reveals incoherent among independently produced forecasts. Each independently produced forecast is subject to different errors, and the summation of these errors from different nodes results in aggregated forecasts that differ from the independently produced forecast at the top level. Consequently, it becomes vital for forecasts to align and aggregate according to the hierarchical structure organizing the array of time series. As a solution to this challenge, forecast reconciliation emerges as one of the most effective methodologies employed to date.

We first consider cross-sectional forecast reconciliation. Recall from Figure **??**, we can construct this hierarchical structure for LTD in a matrix form like this:

$$
\begin{bmatrix}
\text{Total}_t \\
\text{Non-residential}_t \\
\text{Residential}_t \\
\text{Commercial}_t \\
\text{Industrial}_t \\
\text{Other}_t
\end{bmatrix}
=
\begin{bmatrix}
1 & 1 & 1 & 1 \\
0 & 1 & 1 & 1 \\
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
\text{Residential}_t \\
\text{Commercial}_t \\
\text{Industrial}_t \\
\text{Other}_t
\end{bmatrix}
$$

or in a more compact notation:

$$
\mathbf{y}_t = \mathbf{S}\mathbf{b}_t,
$$

where $\mathbf{y}_t$ is an n-dimensional vector of all the observations in the hierarchy at time t, $\mathbf{S}$ represents the summing matrix defining how bottom-level series are aggregated, and $\mathbf{b}_t$ is an m-dimensional vector of all the observations in the bottom level of the hierarchy at time t. In this case, n and m equal to 6 and 4, respectively.

Then for any set of base forecast, denoted as $\mathbf{\hat{y}_h}$, where h is the forecast horizon, all reconciliation forecasting approaches, generating coherent forecasts $\mathbf{\tilde{y}_h}$, can be represented as:

$$
\mathbf{\tilde{y}_h} = \mathbf{S}\mathbf{G}\mathbf{\hat{y}_h},
$$

where $\mathbf{G}$ is a matrix that maps the base forecasts into the bottom level (**hyndman2021forecasting**).

The equation shows that pre-multiplying any set of base forecasts with $\mathbf{SG}$ will return a set of coherent forecasts.

Within the domain of hierarchical time series forecasting, there are three traditional single level approaches for generating forecasts for hierarchical time series. The first, known as the *bottom-up* approach, initiates by producing forecasts for each series at the lowest level and subsequently aggregates these to generate forecasts for the upper levels of the hierarchy. For this approach, $\mathbf{G}$ can be defined as:

$$\mathbf{G} = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

where the first 2 columns zero out the base forecast of the series above the bottom level.

Conversely, the *top-down* approach starts with a forecast at the highest level, which is then disaggregated to lower levels using predetermined proportions—typically based on historical data distributions (**gross1990disaggregation**). For this approach, **G** can be defined as:

$$\mathbf{G} = \begin{bmatrix} p_1 & 0 & 0 & 0 & 0 & 0 \\ p_2 & 0 & 0 & 0 & 0 & 0 \\ p_3 & 0 & 0 & 0 & 0 & 0 \\ p_4 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

where the first column includes the set of proportions that distribute the base forecasts of the top level to the bottom level

Lastly, the *middle-out* approach is a combination of both the bottom-up and top-down methods.

However, the traditional single level approaches may have their limitations since only base forecast at one level is used. In response to this, **wickramasuriya2018optimal** introduced the *MinT* (Minimum Trace) optimal reconciliation methodology, which is based on the finding of a **G** matrix that minimises the total forecast variance within the set of coherent forecasts.

wickramasuriya2018optimal show that the variance-covariance of the h-step-ahead coherent forecast errors is given by:

$$\mathbf{V}_h = Var[\mathbf{y}_{T+h} - \tilde{\mathbf{y}}_\mathbf{h}] = \mathbf{S}\mathbf{G}\mathbf{W_h}\mathbf{G'}\mathbf{S'},$$

where $\mathbf{W}_h = Var[\mathbf{y}_{T+h} - \hat{\mathbf{y}}_\mathbf{h}]$ is the variance-covariance matrix of the corresponding base forecast errors.

wickramasuriya2018optimal also show that:

$$\mathbf{G} = (\mathbf{S}'\mathbf{W}_h^{-1}\mathbf{S})^{-1}\mathbf{S}'\mathbf{W}_h^{-1},$$

minimises the trace of $\mathbf{V}_h$ subject to **S G S = S**

Therefore, the optimally reconciled forecasts are given by:

$$\tilde{\mathbf{y}}_{\mathbf{h}} = \mathbf{S}(\mathbf{S}'\mathbf{W}_{\mathbf{h}}^{-1}\mathbf{S})^{-1}\mathbf{S}'\mathbf{W}_{\mathbf{h}}^{-1}\hat{\mathbf{y}}_{\mathbf{h}}$$

which refers as the **MinT**

Now, a challenge with G is that it requires an estimation of $\mathbf{W}_h$, the forecast error variance of h-step-ahead base forecasts. There are four simplifying approximations in place that have been shown to work well:

1. **OLS** (**hyndman2011optimal**): $\mathbf{W}_h = k_h\mathbf{I}$ for all h, where $k_h > 0$.

This is the most simplifying assumption to make, and means that **G** is independent of the data, providing substantial computational savings. The disadvantage, however, is that this specification does not account for the differences in scale between the levels of the structure, or for relationships between series.

This is the simplest assumption to adopt, implying that **G** is independent of the data, which significantly reduces computational costs. However, this approach does not account for the varying scales across different levels of the structure or for the relationships between series.

2. **WLS$_\mathbf{S}$** (**athanasopoulos2017forecasting**): $\mathbf{W}_h = k_h\mathbf{\Lambda}$ for all h, where $k_h > 0$, $\mathbf{\Lambda} = diag(\mathbf{S1})$, and $\mathbf{1}$ is unit vector of dimension m (the number of bottom-level series).

Applying structural scaling is particularly useful in cases where residuals are not available, and so the variance scaling cannot be applied; for example, in cases where the base forecasts are generated by judgemental forecasting.

3. **WLS$_\mathbf{V}$** (**hyndman2016fast**): $\mathbf{W}_h = k_h\text{diag}(\hat{\mathbf{W}}_1)$ for all h, where $k_h > 0$,

$$\hat{\mathbf{W}}_1 = \frac{1}{T}\sum_{t=1}^{T}\mathbf{e}_t\mathbf{e'}_t,$$

and $\mathbf{e}_t$ is an n-dimensional vector of residuals of the models that generated the base forecasts stacked in the same order as the data.

4. **MinT$_\mathbf{S}$** (**wickramasuriya2018optimal**): $\mathbf{W}_h = k_h\hat{\mathbf{W}}_{1,D}^*$ for all h, where $k_h > 0$, and $\hat{\mathbf{W}}_{1,D}^* = \lambda\hat{\mathbf{W}}_{1,D}^* + (1-\lambda)\hat{\mathbf{W}}_1$ is a shrinkage estimator with diagonal target $\hat{\mathbf{W}}_{1,D}^*$, a diagonal matrix comprising the diagonal entries of $\hat{\mathbf{W}}_1$, and λ the shrinkage intensity parameter.

In contrast to earlier estimators of variance and structural scaling, this method captures strong interrelationships between time series in the hierarchy, while the use of shrinkage helps to manage the complexity of the estimation due to the size of $\mathbf{W}_h$.

By following the above procedure, we will get cross-sectionally coherent forecasts. Moreover, this procedure extends to temporal reconciliation and cross-temporal reconciliation.

Cross-sectional reconciliation

For cross-sectional reconciliation, there are two main challenges, as stated by **kourentzes2019cross**:

1. the size of the cross-sectional dimension of the hierarchy; and
2. the heterogeneity of the series across, but also within levels.

The size relates directly with estimation of $\mathbf{W}_h$, and therefore very large hierarchies, estimation of $\mathbf{W}_h$ can be computationally expensive. However, this is not the case for this hierarchical structure in Figure **??**. On the other hand, it is expected that there will be heterogeneity between each cross-sectional levels. Therefore, in this case, we choose not to apply structural scaling since it assumes that each of the bottom-level base forecasts has errors with equal variance k_h.

Temporal reconciliation

On the other hand, for temporal hierarchies, forecasts across all levels are for the same series, therefore it is safe to assume homogeneity between each level. On the other hand, following the arguments by **athanasopoulos2017forecasting**, since the covariances in $\mathbf{W}_h$ would be between series of different sampling frequencies due to the temporal aggregation, we do not implement the MinT shrinkage estimator.

Cross-temporal reconciliation

The cross-temporal reconciliation methodology used in this context is heuristic first-temporal-then-cross-sectional reconciliation, proposed by **kourentzes2019cross**. It is a 2-step approach, where step 1 includes reconciliation through temporal hierarchies for each single variable to achieve temporally coherence forecast (**THieFs**). Then for step 2, from previously **THieFs**, we generate k cross-sectional reconciliations, setting $\mathbf{W}_h = \hat{\mathbf{W}}_{\mathbf{h},\ell}$, where $\ell = 1, 2, .., k$, and k denotes the number of temporal aggregation levels. This results in reconciliation matrix $\mathbf{SG}_\ell$ for each temporal aggregation level. And by averaging across these, we compose a consensus reconciliation matrix $\mathbf{SG}$, where $\mathbf{G} = \frac{1}{k} \sum_{\ell=1}^{k} \mathbf{G}_\ell$, capturing the reconciliation consensus across all k temporal aggregation levels. The outcome are cross-temporally reconciled forecasts, which are coherent across both dimensions, at all scales.

4.3　Model Selection

As discussed in Section **??**, it highlights the pertinence of employing VECM to adequately model these relationships.

One limitation of these models is their reliance on ample data to produce reliable parameters. Monthly data spanning over a decade for LTD has been found to be sufficient. Nevertheless, the subsequent section on cross-validation will elaborate on the precise sizing of the training dataset required to ensure compatibility with the VAR and VECM models.

In addition to VAR and VECM, the Autoregressive Integrated Moving Average (ARIMA) model has also been employed, primarily for purposes of comparison and validation of improvements. The efficacy of this comparison will be assessed through time series cross-validation, utilizing various accuracy metrics such as:

- *Root Mean Square Error (RMSE)*, where the RMSE is computed as:

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(y_t - \hat{y}_t)^2}; \text{ and}$$

- *Mean Absolute Percentage Error (MAPE)*, where the MAPE is computed by defining the percentage error, to overcome scale-dependency of RMSE as:

$$\text{MAPE} = \frac{1}{n}\sum_{t=1}^{n}\left|100 \times \frac{y_t - \hat{y}_t}{y_t}\right|.$$

5　Time Series Cross-Validation

5.1　Definition and rationale

Time series cross-validation represents a sophisticated adaptation of the conventional training/test set approach for model selection, as stated by **hyndman2021forecasting**. This methodology is particularly well-suited to time series data because it exclusively includes observations from periods prior to those being forecasted, distinguishing it from traditional cross-validation techniques.

In the scope of this project, time series cross-validation is employed to rigorously evaluate whether the VAR/VECM or ARIMA models yield more accurate forecasts for this dataset across forecasting horizons ranging from 1 to 12 steps ahead. Additionally, this approach is used to gauge the extent of improvement introduced by the reconciliation process in comparison to the base forecasts. Time series cross-validation facilitates the generation of a series of 1 to 12-step-ahead forecasts across various rolling periods. For instance, a 12-step-ahead forecast can be produced starting from January 2022,

followed by another 12-step-ahead forecast from February 2022, and so forth. This is essential for the *Department of Treasury and Finance* to analyze the effects of market dynamics or policy changes on land transfer duty.

The initial training set will start with 108 monthly observations, equating to 9 years of data, and will roll forward by 1 month for each subsequent step. This process will generate 10 folds of training sets, and thus there will be 10 sets of 12-step-ahead forecasts with equivalent test set for forecasting accuracy evaluation . While the number of observations in the initial training set can be lower, 108 months were chosen for code efficiency.

5.2   Generating forecast

To enhance forecast accuracy, this methodology extends beyond a cross-sectional hierarchical structure to incorporate a temporal hierarchical structure, culminating in cross-temporal reconciliation forecasting. This approach integrates forecasts across different time structures, and its efficacy is evaluated by comparing it with cross-sectional and temporal reconciliation forecasts individually to determine any improvements.

Given the monthly frequency of the data, additional aggregation levels such as bi-monthly, quarterly, four-monthly, semi-annually, and annually are established, with the annual aggregation representing the top level of the temporal hierarchy.

Models are fitted at each cross-sectional level across these varying temporal frequencies. A loop is utilized to generate forecasts and residuals, which are subsequently organized into a matrix for each temporal frequency: monthly (denoted as k1), bi-monthly (k2), quarterly (k3), four-monthly (k4), semi-annually (k6), and annually (k12). These matrices are then collectively stored within a list.

This structured approach facilitates the fitting of different models tailored to each temporal frequency, allowing for adjustments in argument values as necessary. Moreover, if a distinct model is required for various cross-sectional levels, an `if` condition is employed within each temporal frequency loop to ensure this customization. The forecasts and residuals are then allocated to `base` and `res` data structures, respectively.

ARIMA

The ARIMA model will be implemented across all levels of the temporal hierarchical structure and at every cross-sectional level. Since as discussed in Section **??**, there is seasonal pattern for some nodes, such as total land transfer duty, residential property land transfer duty or sales. And thus we will pay more attention on seasonal ARIMA model. It is a combination of differencing with autoregression and a moving average model. The full model can be written as:

$$y'_t = c + \phi_1 y'_{t-1} + \cdots + \phi_p y'_{t-p} + \Phi_1 y'_{t-m} + \cdots + \Phi_P y'_{t-mP} + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q} + \Theta_1 \varepsilon_{t-m} + \cdots + \Theta_Q \varepsilon_{t-mQ} + \varepsilon_t,$$

where:

- y'_t is the differenced series, considering both non-seasonal and seasonal differencing.
- ϕ_i are the coefficients for the non-seasonal autoregressive (AR) terms.
- θ_j are the coefficients for the non-seasonal moving average (MA) terms.
- Φ_k are the coefficients for the seasonal autoregressive (SAR) terms.
- Θ_l are the coefficients for the seasonal moving average (SMA) terms.
- ε_t is the white noise error term.
- m is the seasonal period.

The parameters are defined as follows:

Parameter	Description
(p)	order of the non-seasonal autoregressive part;
(d)	degree of non-seasonal differencing involved;
(q)	order of the non-seasonal moving average part;
(P)	order of the seasonal autoregressive part;
(D)	degree of seasonal differencing involved;
(Q)	order of the seasonal moving average part;
(m)	number of periods in each season.

Or it can be written as:

$$\text{ARIMA} \quad \underbrace{(p, d, q)}_{\substack{\text{Non-seasonal part} \\ \text{of the model}}} \quad \underbrace{(P, D, Q)_m}_{\substack{\text{Seasonal part} \\ \text{of the model}}},$$

where m is the seasonal period (e.g., number of observations per year)

To implement this, we can use the `auto.arima()` function from the `forecast` package. This function efficiently determines the optimal parameters for the autoregressive lag, differencing, and moving average components of the model by automatically adjusting them for the error terms.

VAR/VECM

A vector autoregression model of order 1, VAR(1) between 2 time series can be represented with:

$$y_t = \beta_{10} + \beta_{11} y_{t-1} + \beta_{12} x_{t-1} + v_t^y$$

$$x_t = \beta_{20} + \beta_{21} y_{t-1} + \beta_{22} x_{t-1} + v_t^x$$

Then the one-step-ahead forecasts are generated by:

$$\hat{y}_{T+1|T} = \hat{\beta}_{10} + \hat{\beta}_{11} y_T + \hat{\beta}_{12} x_T$$

$$\hat{x}_{T+1|T} = \hat{\beta}_{20} + \hat{\beta}_{21} y_T + \hat{\beta}_{22} x_T$$

The VAR(1) equation can be extended to lag of 3 between 3 time series, which is our case:

$$y_t = \beta_{10} + \beta_{11} y_{t-1} + \beta_{12} x_{t-1} + \beta_{13} z_{t-1} + \beta_{14} y_{t-2} + \beta_{15} x_{t-2} + \beta_{16} z_{t-2} + \beta_{17} y_{t-3} + \beta_{18} x_{t-3} + \beta_{19} z_{t-3} + v_t^y$$

$$x_t = \beta_{20} + \beta_{21} y_{t-1} + \beta_{22} x_{t-1} + \beta_{23} z_{t-1} + \beta_{24} y_{t-2} + \beta_{25} x_{t-2} + \beta_{26} z_{t-2} + \beta_{27} y_{t-3} + \beta_{28} x_{t-3} + \beta_{29} z_{t-3} + v_t^x,$$

$$z_t = \beta_{30} + \beta_{31} y_{t-1} + \beta_{32} x_{t-1} + \beta_{33} z_{t-1} + \beta_{34} y_{t-2} + \beta_{35} x_{t-2} + \beta_{36} z_{t-2} + \beta_{37} y_{t-3} + \beta_{38} x_{t-3} + \beta_{39} z_{t-3} + v_t^z$$

where y_t, x_t and z_t represent land transfer duty, sales and home value index.

However, as discussed in Section **??**, given the existence of at least two cointegration relationships among the three time series, which are different type of property land transfer duty against sales and home value index, fitting a VECM is more appropriate. The VECM is formulated by incorporating the error correction term into the VAR model in first differences. For the three time series (y_t), (x_t), and (z_t), the VECM(3) can be expressed as:

$$\Delta y_t = \alpha_1(\gamma_{11} y_{t-1} + \gamma_{12} x_{t-1} + \gamma_{13} z_{t-1} - \mu_1) + \sum_{j=1}^{3} \left(\beta_{1j1} \Delta y_{t-j} + \beta_{1j2} \Delta x_{t-j} + \beta_{1j3} \Delta z_{t-j} \right) + \varepsilon_t^y$$

$$\Delta x_t = \alpha_2(\gamma_{21} y_{t-1} + \gamma_{22} x_{t-1} + \gamma_{23} z_{t-1} - \mu_2) + \sum_{j=1}^{3} \left(\beta_{2j1} \Delta y_{t-j} + \beta_{2j2} \Delta x_{t-j} + \beta_{2j3} \Delta z_{t-j} \right) + \varepsilon_t^x$$

$$\Delta z_t = \alpha_3(\gamma_{31} y_{t-1} + \gamma_{32} x_{t-1} + \gamma_{33} z_{t-1} - \mu_3) + \sum_{j=1}^{3} \left(\beta_{3j1} \Delta y_{t-j} + \beta_{3j2} \Delta x_{t-j} + \beta_{3j3} \Delta z_{t-j} \right) + \varepsilon_t^z$$

where:

- Δy_t, Δx_t, and Δz_t represent the first differences of y_t, x_t, and z_t, respectively.

- α_i are the adjustment coefficients for the error correction terms.
- γ_{ij} are the coefficients of the cointegrating vectors.
- μ_i are the constants in the cointegration relationships.
- ε_t^y, ε_t^x, and ε_t^z are the white noise error terms.

The VECM framework thus captures both the short-term dynamics through the differenced terms and the long-term equilibrium relationships through the error correction terms.

Finally, the reconciliation procedure can be implemented using the `FoReco` package **FoReco**.

6 Results

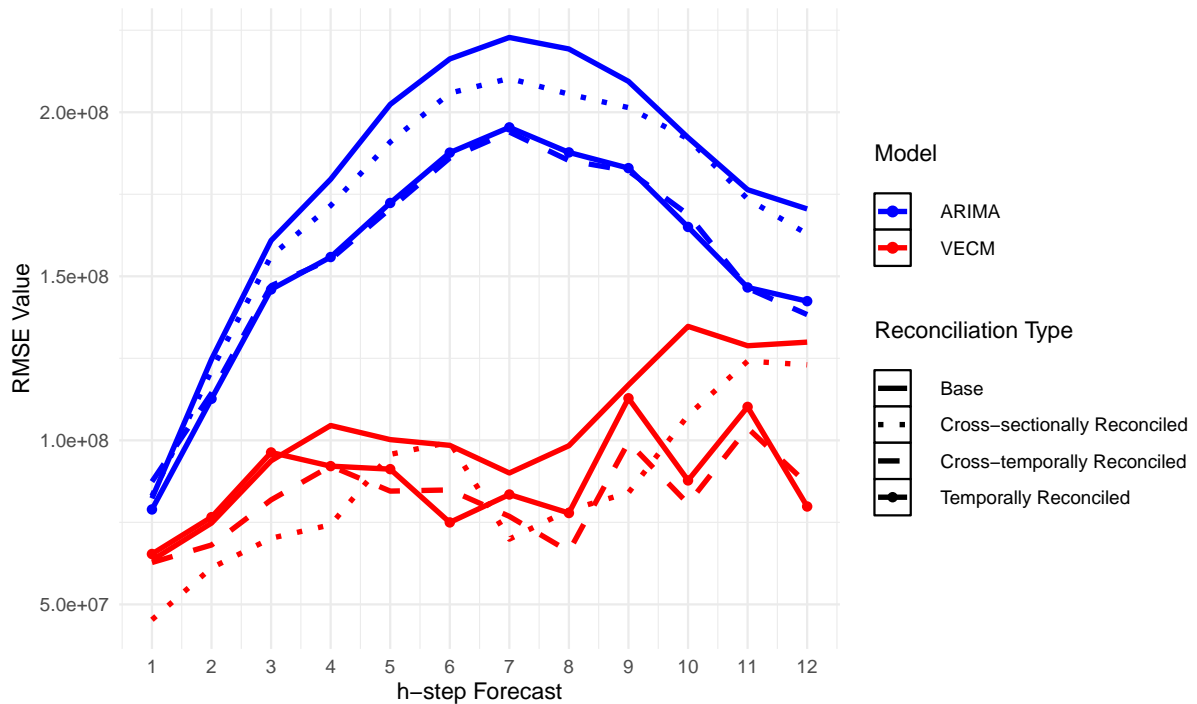## 6.1 RMSE

Figure 27: *Monthly total LTD base and reconciled forecast RMSE by model*

Figure **??** illustrates the Root Mean Square Error (RMSE) of reconciled forecasts in comparison to base forecasts for both VECM and ARIMA models. The blue lines denote the ARIMA model, while the red lines represent the VECM. The line types are differentiated as follows: solid lines for base forecasts, solid lines with circle markers for temporally reconciled forecasts, dotted lines for cross-sectionally reconciled forecasts, and dashed lines for cross-temporally reconciled forecasts.

Upon initial examination, it is evident that the VECM model consistently produces more accurate forecasts compared to the ARIMA model, aligning with the model selection by the *Department of*

Treasury and Finance. Furthermore, the significant disparity between reconciled forecasts and base forecasts, particularly from the four-step-ahead forecasts onwards for both models, suggests that forecast reconciliation markedly enhances forecast accuracy. Notably, for both models, the RMSE lines of temporally reconciled forecasts are closer to those of cross-temporally reconciled forecasts, compared to the RMSE lines of cross-sectionally reconciled forecasts. This indicates that temporal hierarchies play a more substantial role in improving forecast accuracy compared to cross-sectional hierarchies.

It is also noteworthy that for one to three-step-ahead forecasts, the base forecasts already exhibit high performance, hence the improvements from applying forecast reconciliation are not evident.

6.2 MAPE

Using the Mean Absolute Percentage Error (MAPE), we anticipate results consistent with those obtained from the Root Mean Square Error (RMSE), which is indeed the case as demonstrated in Figure Figure **??**.
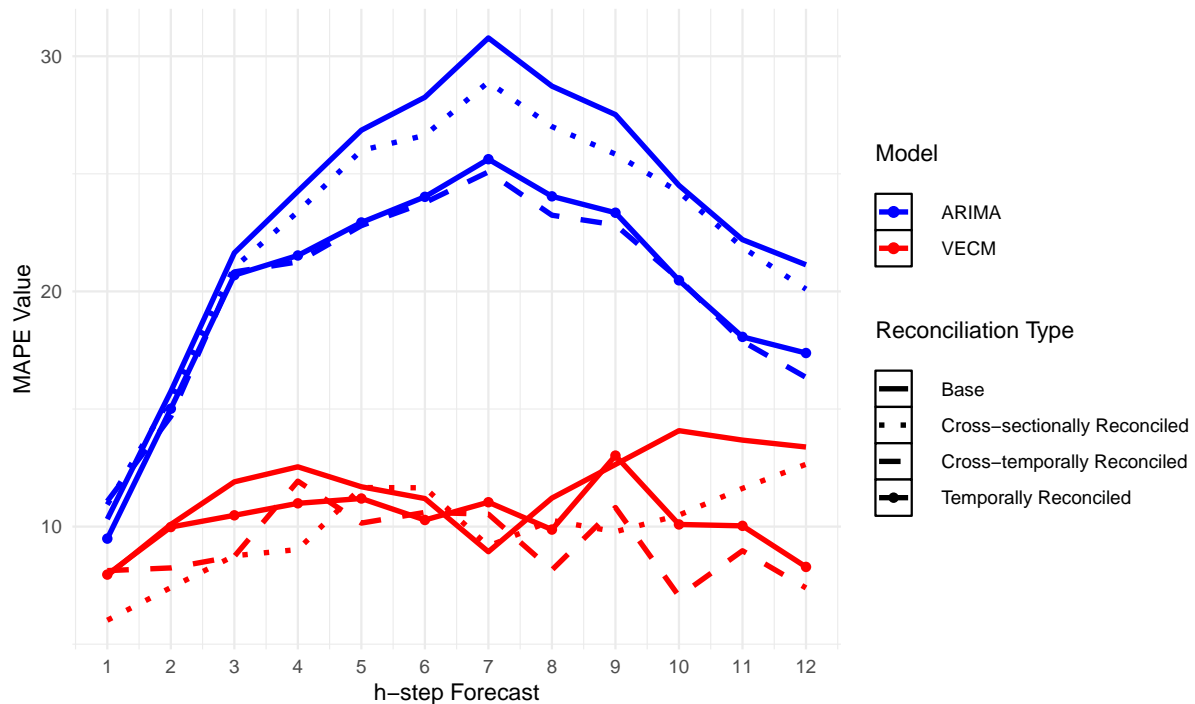

Figure 28: *Monthly total LTD base and reconciled forecast MAPE by Method*

In conclusion, four key insights can be derived from Figure **??** and Figure **??**:

1. The VECM model is more suitable than the ARIMA model.

2. Forecast reconciliation significantly enhances forecast accuracy, with cross-temporally reconciled forecasts exhibiting the best performance.

3. The majority of the improvement is attributable to incorporating temporal hierarchies.

4. Forecast reconciliation is particularly effective in improving forecasts that initially perform poorly.

6.3 Comparing against DTF forecast

The Root Mean Squared Errors (RMSE) of the *Department of Treasury and Finance* (DTF) forecast were computed utilizing the same time-series cross-validation procedure as previously described, with the original scale of land transfer duty serving as a benchmark. As shown in Figure **??**, the RMSE varies between 110 million and 130 million, with a significant reduction observed at the 12-step-ahead forecasts. This reduction may be attributed to a strong seasonal pattern at step 12, which is overlooked when using the original scale of land transfer duty. Overall, the RMSE derived from the DTF forecast is not superior to the RMSE of the reconciled forecasts.
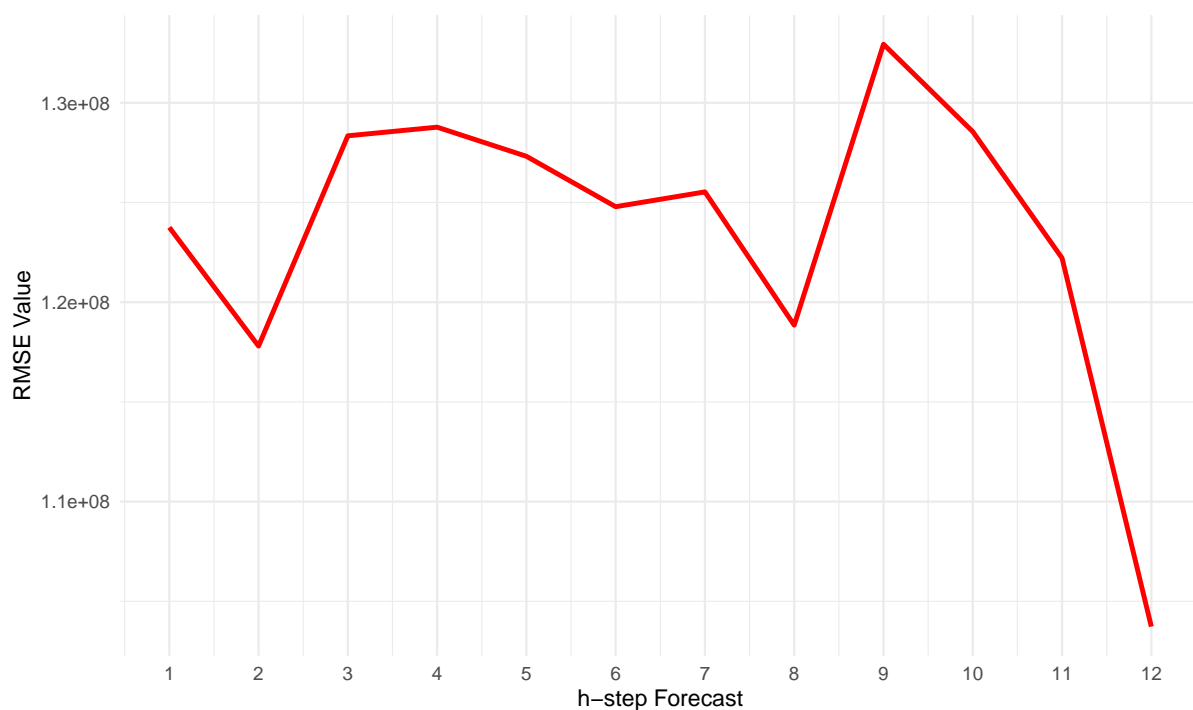

Figure 29: *RMSE of DTF forecast*

When extending the comparison to cross-temporally reconciled forecasts against DTF forecasts, as illustrated in Figure **??**, where the red line represents the ratio of MAPE of reconciled forecasts to DTF forecasts, it is evident that cross-temporally reconciled forecasts perform significantly better.
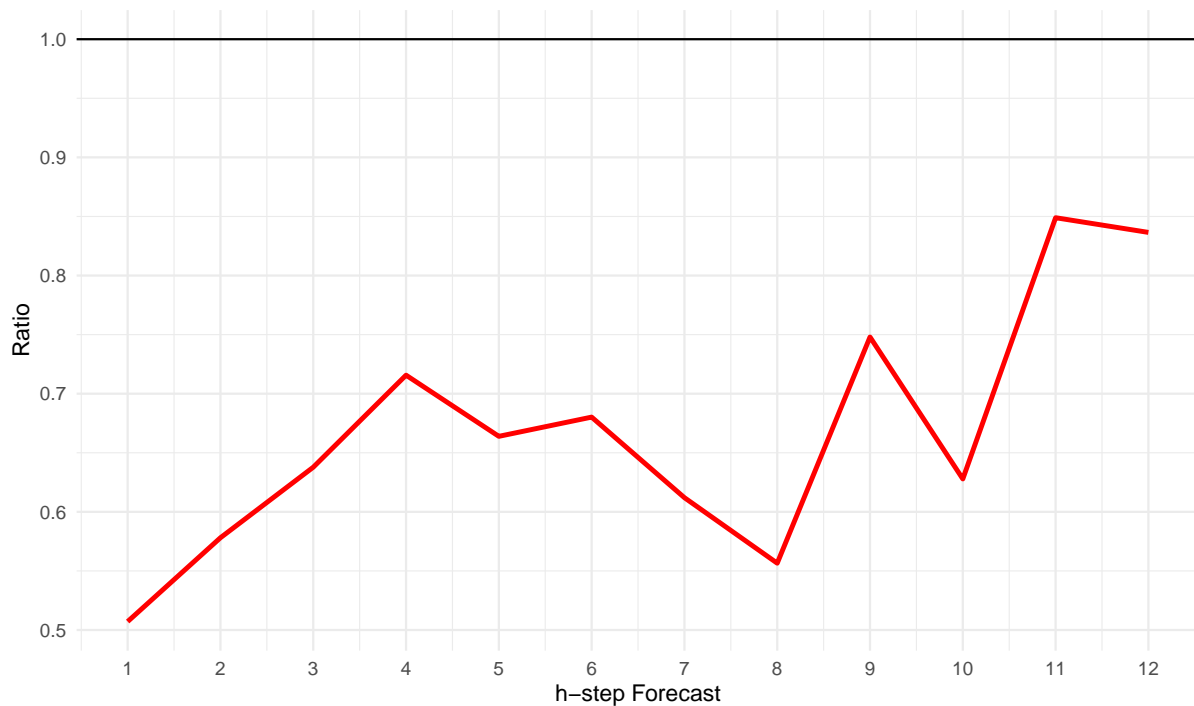
7 Discussion

Figure 30: *Ratio of RMSE of cross-temporally reconciled forecast to DTF forecast*

7.1 Interpretation of Results

The empirical analysis reveals that the VECM model provides more accurate forecasts for LTD compared to the ARIMA model, particularly when forecast reconciliation techniques are applied. The incorporation of cross-sectional and temporal hierarchies into a unified cross-temporal reconciliation framework markedly improves forecast accuracy. This indicates that the proposed methodology effectively captures both short-term dynamics and long-term trends in the property market.

Temporal reconciliation emerges as more impactful than cross-sectional reconciliation in enhancing forecast accuracy, underscoring the importance of temporal aggregation in the forecasting process.

The comparative analysis with the DTF's forecasts further validates the proposed methodology, demonstrating that cross-temporally reconciled forecasts consistently outperform the DTF's forecasts, as evidenced by lower RMSE values.

7.2 Limitations

- **Seasonal Adjustment**: Although there is a discernible seasonal pattern in land transfer duty (LTD), residential LTD (`ltd_res`), and sales, we did not apply seasonal adjustments. Ignoring seasonal adjustments might have affected the model's ability to accurately capture seasonal variations, potentially impacting the forecast accuracy. Future research could incorporate seasonal adjustment methods to address this limitation.

- **Non-Stationarity in Cointegration Analysis**: When fitting the VECM with sales and Home Value Index (HVI) to LTD for non-residential, commercial, industrial, and other property types, the Augmented Dickey-Fuller (ADF) test on the linear combination of these series, based on the eigenvalues from the Johansen test, indicated non-stationarity. Non-stationary time series can lead to unreliable model estimates and forecasts. To improve stationarity, alternative variables could be explored to replace sales and HVI.

- **Lag Selection for Johansen Test**: The lag length for the Johansen test during the 10-fold cross-validation was selected via a manual trial-and-error procedure, aiming to achieve the lowest RMSE. This manual approach may not identify the optimal lag length consistently. Future improvements could involve using cross-validation or other automated procedures to systematically determine the optimal lag length, enhancing the robustness and reliability of the model selection process.

8   Conclusion and Recommendations

This study introduced an advanced forecasting methodology for land transfer duty (LTD) in Victoria, leveraging Vector Error Correction Model (VECM) and Autoregressive Integrated Moving Average (ARIMA) models, enhanced by cross-sectional, temporal, and cross-temporal hierarchies. The findings underscore the superiority of VECM, especially when coupled with forecast reconciliation techniques, in producing more accurate forecasts compared to ARIMA. Temporal hierarchies were found to be more influential than cross-sectional hierarchies in improving forecast accuracy. The reconciled forecasts consistently outperformed those of the *Department of Treasury and Finance*, as evidenced by lower RMSE values. Despite some limitations, such as the lack of seasonal adjustment and manual lag selection for the Johansen test, the methodology demonstrates significant potential for enhancing LTD forecasting accuracy. Future work should focus on addressing these limitations to further refine the approach and improve its applicability.

9   GitHub repo

Here is the link to the GitHub repo of this project.