Modelling Seasonality in Australian Building Approvals

Harry Karamujic, (The University of Melbourne, Australia)

Abstract
The paper examines the impact of seasonal influences on Australian housing approvals, represented by the State of Victoria building approvals for new houses (BANHs). The prime objective of BANHs is to provide timely estimates of future residential building work. Due to the relevance of the residential property sector to the property sector as whole, BANHs are viewed by economic analysts and commentators as a leading indicator of property sector investment and as such the general level of economic activity and employment. The generic objective of the study is to enhance the practice of modelling housing variables. In particular, the study seeks to cast some additional light on modelling the seasonal behaviour of BANHs by: (i) establishing the presence, or otherwise, of seasonality in Victorian BANHs; (ii) if present, ascertaining is it deterministic or stochastic; and (iii) determining out of sample forecasting capabilities of the considered modelling specifications. To do so the study utilises a structural time series model of Harvey (1989). The modelling results confirm that the modelling specification allowing for stochastic trend and deterministic seasonality performs best in terms of diagnostic tests and goodness of fit measures. This is corroborated with the analysis of out of sample forecasting capabilities of the considered modelling specifications, which showed that the models with deterministic seasonal specification exhibit superior forecasting capabilities. The paper also demonstrates that if time series are characterized by either stochastic trend or seasonality, the conventional modelling approach (a modelling approach based on the assumption of deterministic trend and deterministic seasonality) is bound to be mis-specified i.e. would not be able to identify statistically significant seasonality in time series. According to the selected modeling specification, factors corresponding to June, April, December and November are found to be significant at five per cent level.

Keywords: New housing building approvals, Univariate structural time series modelling, Out of sample forecasting, Stochastic and deterministic trend, Stochastic and deterministic seasonality

Introduction
Building approvals for new houses (BANHs) denotes the number of new house building works approved. From July 1990, the statistics include all approved new residential building valued at A$10,000 or more. A new house can be defined as the construction of a detached building that is primarily used for long term residential purposes. According to the Australian Bureau of Statistics (ABS) (2009), statistics of building work approved are compiled from “permits issued by local government authorities and other principal certifying authorities, and contracts let or day labour work authorised by commonwealth, state, semi-government and local government authorities”.

As with most other countries worldwide, all houses built in Australia have to be approved by relevant government departments before building commences. Consequently, BANHs are used to inform as to how many new buildings are expected to be constructed in the near future. Because BANHs provide timely estimates of future residential building work, and due to the relevance of the residential property sector to the property sector as whole, BANHs are viewed by economic analysts and commentators as a leading indicator of property sector investment and as such the general level of economic activity and employment.
The awareness of the sizeable impact of the property sector on the soundness of financial institutions and the level of economic activity is not a new observation. It is commonly accepted that the boom-bust nature of property prices plays an important role in explaining business cycles (by strengthening the upswing and amplify the downswing). Typically, reducing housing prices tend to impose additional pressure on the banking sector. This happens not only because of increases in bad debts for mortgage loans, but also because of the deterioration in the balance sheets of corporate borrowers that rely on property as collateral. Not surprisingly, fluctuations in housing activity and the extent to which they interact with the financial sector and the whole economy are very much of interest, among others, to the government, the reserve bank and other financial regulators.

The mainstream literature have recognised for a long time that investment in housing and consumer durables lead non-residential business fixed investment over the business cycle (e.g. Burns and Mitchell 1946). Among others, this is corroborated by Fisher (2006), who observed that in seven of the ten post-war recessions in the USA, household investment achieved its peak and trough before business investment. Ball and Wood (1999) conducted comparative structural time series analysis of housing investment in advanced world economies. They looked at the impact of housing investment on the economy and concluded that housing investment fluctuations after 1960s become a destabilizing factor. This finding highlighted the significance of this category of investment and further accentuated the relevance of studies focusing on better understanding housing investment volatility.

This study examines the impact of seasonal influences on Australian housing approvals, represented by Victorian BANHs. Victoria has been selected as a test case because of its geographical homogeneity and economic relevance. Victoria is Australia's most urbanized state: nearly 90 per cent of residents living in cities and towns, it is the most densely populated state (22 people on square kilometer), and has a highly centralized population, with almost 75 per cent of Victorians living in the state capital and largest city, Melbourne. At the same time, the state of Victoria is the second largest economy in Australia, after New South Wales, accounting for almost a quarter of the nation's gross domestic product (GDP). According to the ABS (2011), in 2008/2009 Victoria contributed 22.6 per cent of the Australian GDP. All other Australian states are either geographically dispersed (cover a wide geographical region across different time and climatic zones) or economically much less significant.

It is important to note that the focus of the study is not on modeling the behaviour of time series in terms of explanatory variables (the conventional modeling approach). The conventional modeling approach assumes that the behaviour of the trend and seasonality can be effectively captured by a conventional regression equation that assumes deterministic trend and seasonality. Instead, the aim is to use a univariate structural time series modeling approach (allows modeling both stochastic and deterministic trend and seasonality) and show that conventional assumptions of deterministic trend and seasonality are not always applicable. Within this specification all components are stochastic; nevertheless each can turn deterministic as a limiting case. In addition to enhancing the practice of modelling an important housing variable, the objective of this paper to empirically assesses the presence of seasonal variations in Victorian housing approvals. Specifically, the study seeks to cast some additional light on BANHs by: (i) establishing the presence, or otherwise, of seasonality in Victorian BANHs, (ii) if present, ascertaining if it is deterministic or stochastic, and (iii) determining out of sample forecasting capabilities of the considered modeling specifications. To do so the study utilises the basic structural time series model of Harvey (1989). Compared to the conventional procedure, Harvey’s (1989) structural time series model involves an explicit modelling of seasonality as an unobserved component.

Empirical evidence of seasonal variations in property related variables is relatively limited. Studies come from a range of different perspectives and employ a number of modelling
techniques. Harris (1989) provided empirical evidence of strong second and third quarter seasonality in the USA house prices. Ma and Goebel (1991) established the presence of January seasonal effect for securitised mortgage markets, while Friday and Peterson (1997) and Colwell and Park (1990) established presence of a January seasonal effect in returns of Real Estate Investment Trusts (REITs) in the USA. Rossini (2000) examined seasonal effects in the housing markets of Adelaide, South Australia, and, with respect to the volume of detached dwelling transactions, determined the presence of statistically significant ‘summer’ and ‘autumn’ seasonal effects. Similarly, Costello (2001) examined the impact of seasonal influences on housing market activity in Perth, Western Australia, and found that the volume of transactions and hence demand is greatest during the first quarter of a year and lowest during the last quarter. Karamujic (2009) confirmed the presence of both cyclical and seasonality in Australian residential mortgage interest rates in the two major Australian banks (National Australia Bank (NAB) and Commonwealth Australia Bank (CBA)). All studies, to varying degrees, point to the existence of seasonality. A study of the literature on housing variables uncovered that most studies are focusing on house prices and found no empirical research focusing on seasonal fluctuations in BANHs.

The rest of the paper is organised as follows. The next section of this paper outlines the methodology used. Section 3 elaborates on data specification, modelling test results and interpretation of the modelling results. Finally, in Section 4, the paper concludes.

Isolating the Seasonal Component: Methodology

Modelling a changing seasonal component is relatively easy for quarterly and monthly observations, with the seasonal component normally being combined with a stochastic trend and an irregular term. This is either done explicitly, as in the structural time series modelling approach, or implicitly, as in the integrated moving average (ARIMA) approach. In the latter case, the seasonal component is specified by means of a canonical decomposition as shown by Hillmer and Tao (1982). The seasonal component can be extracted by a state space smoothing algorithm; see for example, Kitagawa and Gersch (1984) or Harvey (1989).

Carrying out such model-based seasonal adjustment, using either approach, has considerable attractions because the procedure adapts to the particular characteristics of the series involved.

A structural time series framework approach, used in this paper, is in line with that promulgated by Harvey (1989). Such models can be interpreted as regressions on functions of time in which the parameters are time-varying. This makes them a natural vehicle for handling changing seasonality of a complex form. Once a suitable model has been fitted, the seasonal component can be extracted by a smoothing algorithm.

Following Harvey (1989) and Harvey, Koopman and Riani (1997), the basic structural time series model is formulated in terms of a trend, seasonal and irregular components. All are assumed to be stochastic and driven by serially independent Gaussian disturbances that are mutually independent. If there are $s$ seasons in the year, the model is

$$y_t = \mu_t + \gamma_t + \varepsilon_t, \quad \varepsilon_t \sim NID(0, \sigma^2),$$  

where the trend, seasonal and irregular are denoted by $\mu_t$, $\gamma_t$ and $\varepsilon_t$, respectively. The trend is specified as follows:

$$\mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t, \quad \eta_t \sim NID(0, \sigma^2),$$  

$$\beta_t = \beta_{t-1} + \zeta, \quad \zeta_t \sim NID(0, \sigma^2)$$
where \( \mu \) is the level and \( \beta \) is the slope. The disturbances \( \eta_t \) and \( \zeta_t \) are assumed to be mutually independent. Setting \( \sigma^2_\eta = 0 \) gives a trend that is relatively smooth.

The seasonal component is generally constructed in terms of stochastic trigonometric functions at the \( s/2 \) seasonal frequencies, although deterministic and dummy-variable formulations are also possible. The fundamental point is that, although the seasonal component is non-stationary, it has the property that the expected value of the sum over the previous \( s \) time periods is zero. This ensures that the seasonal effects are not confounded with the trend. It also means that the forecasts of the seasonal component will sum to zero over any one-year period. The statistical treatment of the model is based on the state-space form, with \( s + 1 \) elements in the state vector. Estimation, forecasting and signal extraction are carried out by means of the Kalman filter and associated algorithms.

The trigonometric form of stochastic seasonality used in models of the form (1) where \( s \) seasons in the year is

\[
\gamma_t = \sum_{j=1}^{[s/2]} \gamma_{j,t}, \quad t = 1, \ldots, T, \tag{3}
\]

and each \( \gamma_{j,t} \) is generated by

\[
\begin{bmatrix}
\gamma_{j,t} \\
\gamma_{j,t}'
\end{bmatrix} = \begin{bmatrix}
\cos \lambda_j & \sin \lambda_j \\
-\sin \lambda_j & \cos \lambda_j
\end{bmatrix} \begin{bmatrix}
\gamma_{j-1,t}' \\
\gamma_{j-1,t}
\end{bmatrix} + \begin{bmatrix}
\omega_{j,t} \\
\omega_{j,t}'
\end{bmatrix}, \tag{4}
\]

where \( \lambda_j = 2\pi j / s \) is frequency, in radians, for \( j = 1, \ldots, [s/2] \) and \( \omega, \omega' \) are two mutually uncorrelated white-noise disturbances with zero means and common variance \( \sigma^2_\omega \). The basic structural model consisting of the stochastic trend in (2) with trigonometric seasonality is easily put in state-space form by defining the \( (s+1) \times 1 \) state vector \( \alpha_t = (\mu_t, \beta_t, \gamma'_{1,t}, \gamma'_{2,t}, \gamma_{2,t}, \ldots) \). The measurement equation is then

\[
y_t = (1,0,z_t)\alpha_t + \epsilon_t, \tag{5}
\]

where \( z_t = (1,0,1,0,\ldots) \). If the Kalman filter is initiated with a diffuse prior, as shown by De Jong (1991), an estimator of the state with a proper prior is effectively constructed from the first \( s + 1 \) observations.

On the other hand, if we choose to fix the seasonal pattern in (1), thus specifying a deterministic seasonal component, \( \gamma_t \) may be modeled as:

\[
\gamma_t = \sum_{j=1}^{s} \gamma_j z_{jt} \tag{6}
\]

Where \( s \) is the number of seasons and the dummy variable \( z_{jt} \) is one in season \( j \) and zero otherwise. In order not to confound trend with seasonality, the coefficients, \( \gamma_j \), \( j = 1, \ldots, s \), are constrained to sum to zero. The seasonal pattern may be allowed to change
over time by letting the coefficients evolve as random walks as in Harrison and Stevens (1976). If \( \gamma_j \) denotes the effect of season \( j \) at time \( t \), then

\[
\gamma_{jt} = \gamma_{j,t-1} + \omega_j, \quad \omega_j \sim \text{NID}(0, \sigma^2) \quad j = 1, \ldots, s.
\]

Although all \( s \) seasonal components are continually evolving, only one affects the observations at any particular point in time, that is \( \gamma_t = \gamma_{jt} \) when season \( j \) is prevailing at time \( t \). The requirement that the seasonal components evolve in such a way that they always sum to zero is enforced by the restriction that the disturbances sum to zero at each point in time. This restriction is implemented by the correlation structure in

\[
\text{Var}(\omega) = \sigma^2_s (I - s^{-1}ii')
\]

where \( \omega = (\omega_1, \ldots, \omega_s)' \), coupled with initial conditions requiring that the seasonal sum to zero at \( t = 0 \). It can be seen from the equation above that \( \text{Var}(i'\omega) = 0 \).

In the basic structural model, \( \mu_t \) in (1) is the local linear trend of (2), the irregular component, \( \epsilon_t \), is assumed to be random, and the disturbances in all three components are taken to be mutually uncorrelated. The signal-noise ratio associated with the seasonal, that is \( q_s = \sigma^2_s / \sigma^2_\epsilon \), determines how rapidly the seasonal changes relative to the irregular. An example of how the basic structural model successfully captures changing seasonality can be found in the study of alcoholic beverages by Lenten and Moosa (1999).

**Results and Discussion**
The structural time series model represented by (1) is applied to seasonally unadjusted monthly BANHs data for Victoria. All three considered modelling specifications include trend (composed of the level \( (\mu_t) \) and slope \( (\beta_t) \)), seasonal \( (\gamma_t) \) and irregular components \( (\epsilon_t) \). An irregular component is also known as the standard error, which can be defined as the standard deviation of the sampling distribution of the estimator, calculated as the square root of the one-step ahead prediction error variance. The data has been sourced from the ABS. The observed period was from 2000:06 to 2009:05. For consistency, the sample for each variable is standardised to start with the first available June observation and end with the latest available May observation, i.e. \( \gamma_{t-1} \) is a seasonal factor that relates to the last month in the sample (May), \( \gamma_{t-2} \) corresponds to April, \( \gamma_{t-3} \) corresponds to March, and so on. Table 1 also reports on the goodness of fit and diagnostics tests. The goodness of fit measures entail the coefficient of determination (R2), a modified coefficient of determination calculated on the basis of the seasonal mean (Rs2), Akaike’s Information Criterion (AIC), and the Schwarz Bayesian Criterion (BIC). On the other hand, diagnostic tests cover serial correlation, normality and heteroscedasticity. They are represented by the Durbin-Watson statistic (DW), the Bowman-Shenton (1975) test for normality of the residual (N), the Ljung-Box (1978) test for serial correlation (Q) and a test for heteroscedasticity (H).

As shown in Table 1, the analysis considers three modelling specifications. Two modelling specifications (Models one and two) include a stochastic trend, while Model three incorporates a deterministic trend. On the other hand, Model one incorporates stochastic seasonal component, while Models two and three incorporate deterministic (fixed) seasonal
components. In summary, Table 1 reports the estimated components of the state vector \((\mu, \beta, \gamma_{1-11}, \varepsilon_i)\), their t-statistics, goodness of fit measures and diagnostics test statistics.

With respect to the goodness of fit, all assessed models are well defined. Overall, the diagnostic tests are also predominately passed. The only exception is the test for serial correlation (Q), for the Model two (which is slightly above the statistically acceptable level) and Model three (significantly above the statistically acceptable level). The Q statistic for Model three indicates that the model suffers from serial correlation, implying a mis-specified model. In all cases the slope is insignificant and the level is significant.

<table>
<thead>
<tr>
<th>State Variable/Test Statistic</th>
<th>Model 1 (Stochastic Trend and Stochastic Seasonality)</th>
<th>Model 2 (Stochastic Trend and Deterministic Seasonality)</th>
<th>Model 3 (Deterministic Trend and Deterministic Seasonality)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\mu_i)</td>
<td>2533.30 (17.28)</td>
<td>2533.30 (17.28)</td>
<td>2530.80 (34.04)</td>
</tr>
<tr>
<td>(\beta_i)</td>
<td>1.17 (0.05)</td>
<td>1.17 (0.05)</td>
<td>-0.74 (-0.61)</td>
</tr>
<tr>
<td>(\gamma_1)</td>
<td>159.39 (3.87)</td>
<td>288.62 (3.53)</td>
<td>244.62 (1.98)</td>
</tr>
<tr>
<td>(\gamma_2)</td>
<td>242.47 (5.79)</td>
<td>154.24 (1.89)</td>
<td>119.13 (0.96)</td>
</tr>
<tr>
<td>(\gamma_3)</td>
<td>47.67 (1.39)</td>
<td>266.56 (3.28)</td>
<td>240.09 (1.94)</td>
</tr>
<tr>
<td>(\gamma_4)</td>
<td>-97.10 (-2.82)</td>
<td>31.38 (0.39)</td>
<td>13.27 (0.11)</td>
</tr>
<tr>
<td>(\gamma_5)</td>
<td>-113.7 (-3.45)</td>
<td>94.81 (1.17)</td>
<td>84.78 (0.69)</td>
</tr>
<tr>
<td>(\gamma_6)</td>
<td>89.34 (2.71)</td>
<td>100.83 (1.25)</td>
<td>98.83 (0.80)</td>
</tr>
<tr>
<td>(\gamma_7)</td>
<td>134.15 (4.13)</td>
<td>-390.32 (-4.82)</td>
<td>-384.97 (-3.12)</td>
</tr>
<tr>
<td>(\gamma_8)</td>
<td>63.14 (1.95)</td>
<td>-657.19 (-8.12)</td>
<td>-644.56 (-5.22)</td>
</tr>
<tr>
<td>(\gamma_9)</td>
<td>13.54 (0.42)</td>
<td>-50.03 (-0.62)</td>
<td>-30.38 (-0.25)</td>
</tr>
<tr>
<td>(\gamma_{10})</td>
<td>5.16 (0.16)</td>
<td>20.75 (0.26)</td>
<td>47.13 (0.38)</td>
</tr>
<tr>
<td>(\gamma_{11})</td>
<td>-21.80 (-0.96)</td>
<td>-78.88 (-0.97)</td>
<td>-46.02 (-0.37)</td>
</tr>
<tr>
<td>(\varepsilon_i)</td>
<td>277.23</td>
<td>274.39</td>
<td>364.48</td>
</tr>
<tr>
<td>(R^2)</td>
<td>0.30</td>
<td>0.31</td>
<td>-0.21</td>
</tr>
<tr>
<td>(R^2_d)</td>
<td>0.58</td>
<td>0.59</td>
<td>0.33</td>
</tr>
<tr>
<td>AIC</td>
<td>11.55</td>
<td>11.51</td>
<td>12.04</td>
</tr>
<tr>
<td>BIC</td>
<td>11.94</td>
<td>11.88</td>
<td>12.36</td>
</tr>
<tr>
<td>DW</td>
<td>2.09</td>
<td>2.09</td>
<td>0.80</td>
</tr>
<tr>
<td>N</td>
<td>3.45</td>
<td>6.37</td>
<td>1.791</td>
</tr>
<tr>
<td>Q</td>
<td>9.64</td>
<td>15.08</td>
<td>125.71</td>
</tr>
<tr>
<td>H</td>
<td>0.37</td>
<td>0.41</td>
<td>0.29</td>
</tr>
</tbody>
</table>

Table 1 Estimated Coefficients of Final State Vector
As shown in Table 1, out of the three assessed modelling specifications, Model two has the highest $R^2_s$ and the lowest $\varepsilon_t$. On the other hand, Model three (characterised by deterministic trend and deterministic seasonality) with negative $R^2_s$ implies that the model is badly determined i.e. the model is worse than a seasonal random walk model. Overall, all of goodness of fit measures infer that Model three is significantly inferior to Models one and two, and that Model two is somewhat better than Model one.

Figures 1, 2 and 3 provide a visual interpretation of the seasonal elements for each considered modelling specification. The seasonal components evidenced in each of the figures show a constant repetitive pattern over the sample period, providing a visual evidence of the deterministic nature of the seasonal component (fixed seasonal components) in the number of new dwellings approved in Victoria. Figure 4 shows this even more clearly with individual monthly seasonals represented by horizontal lines, implying an unchanging seasonal effect across the whole sample period.

![Figure 1 Model 1 - Seasonal Component](image1)

Out of the eleven seasonal factors relating to the Model two, presented in Table 1, factors corresponding to June ($\gamma_1$), April ($\gamma_3$), December ($\gamma_7$) and November ($\gamma_8$) are found to be significant at five per cent level (shown as variables with t statistics values above 1.96). The factors corresponding to December ($\gamma_7$) and November ($\gamma_8$) exhibit the season-related reduction in the number of BANHs, while the factors corresponding to June ($\gamma_1$) and April ($\gamma_3$) demonstrate season-related increases.

![Figure 2 Model 2 - Seasonal Component](image2)
In summary, the analysis points out that the behaviour of BANHs exhibits stochastic trend and deterministic seasonality. As a result, any model based on assumptions of deterministic trend and seasonality is bound to be mis-specified.
modeling specifications characterised by the fixed seasonal factors, supporting the earlier finding that the seasonal pattern in the number of dwelling units approved in Victoria is deterministic and not stochastic.

<table>
<thead>
<tr>
<th></th>
<th>Models 1 and 2</th>
<th>Model 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sum of absolute errors</td>
<td>9,987</td>
<td>27,574</td>
</tr>
<tr>
<td>Sum of squared errors</td>
<td>5,063,177</td>
<td>19,201,467</td>
</tr>
</tbody>
</table>

Table 2 Sum of absolute / squared errors of the forecasting values

![Figure 5 Out-of-sample Forecasting](image)

Conclusion
The study uses Harvey’s (1989) univariate structural time series mode to examine the impact of seasonal influences on the Australian housing time series, with the generic objective of enhancing the practice of modelling housing variables. Specifically, the paper seeks to cast some additional light on the seasonal behaviour of BANHs by: (i) establishing the presence, or otherwise, of seasonality in Victorian BANHs, (ii) if present, ascertaining is it deterministic or stochastic, and (iii) determining out of sample forecasting capabilities of the considered models.

This is done by estimating three modelling specifications comprised of stochastic and deterministic trend and seasonal components. The goodness of fit measures and the diagnostic test statistics indicate that Model two, which is comprised out of stochastic trend and deterministic seasonality, is unambiguously superior to the other two specifications.

The examination of the out-of-sample forecasting power of the three models clearly shows that the seasonality apparent in the actual data is well picked up by specifications entailing deterministic seasonal factor, corroborating the earlier finding that the seasonal pattern in the number of dwelling units approved in Victoria is deterministic and not stochastic. Evidently the analysis of the three presented modelling specifications indicates that the conventional modelling approach, characterised by assumptions of deterministic trend and deterministic seasonality, would not identify seasonal behaviour of time series characterised by either stochastic trend or seasonality.
The analysis of Model two points out that the behaviour of BANHs exhibits statistically significant seasonal components. A possible explanation for the observed statistically significant reduction in BANHs during December ($\gamma_1$) and November ($\gamma_8$) is the reduction of the level of activity caused by approaching to the ‘summer holidays’ season, while the season-related increases during June ($\gamma_1$) and April ($\gamma_3$) may be explained by a spike in the level of activity during the ‘end of financial year’ season and preparation for a surge in contraction activity during the ‘spring’ season.

To corroborate the modelling results and explanations provided, the scope of the analysis would need to be extended. It is reasonable to expect that these substantial season-related changes in monthly BANHs are, to a large extent, correlated with home loan drawdowns, housing starts and house prices. Thus, extending the research to include home loan drawdowns and housing starts and house prices could be a rewarding area for further research.

References


