Forecasting land transfer duty with a regime switching vector error correction model

By Sebastian Paz and Jonathan Dark[1]

Department of Treasury and Finance

ABSTRACT

Accurately forecasting government revenue is crucial for effective budgeting and policy making. To model and forecast land transfer duty we use regime-switching vector error correction models. We identify two states: a high-growth low-volatility state and a low-growth high-volatility state. In real time, the models identify the high-volatility state associated with COVID-19. Our models consistently provide better forecasts than state invariant alternatives across multiple horizons and loss functions. Our findings suggest that regime switching models may be a useful tool for budgeting purposes.

1. Introduction

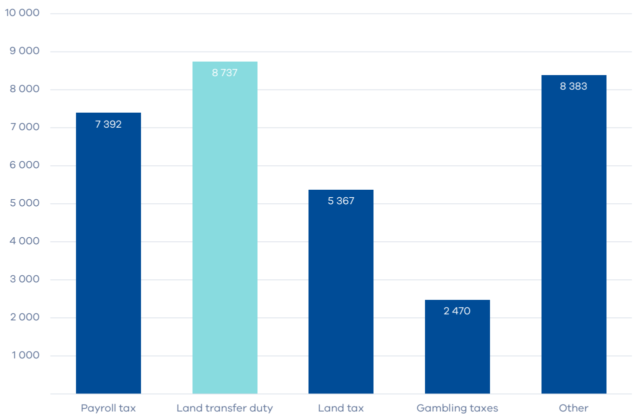
Land transfer duty (LTD), commonly referred to as stamp duty, is a tax payable by property purchasers in Victoria, Australia. LTD rates range from 1.4 per cent to 6.5 per cent of the property's purchase price, depending on its value (State Revenue Office Victoria, 2024). As of 2023, the total value of the Victorian residential housing market is approximately AUD $2 trillion, with a median house price in Melbourne of $907 000. Housing prices have seen an annualised growth rate of around 5 per cent in recent years (CoreLogic Australia, 2023).

LTD accounts for 27 per cent of the State's total taxation revenue in 2023, as shown in Figure 1a. Moreover, LTD returns exhibit substantial volatility – almost five times greater than that of the ASX 200 index – with month-to-month changes ranging from -53.3 per cent to +81.5 per cent.

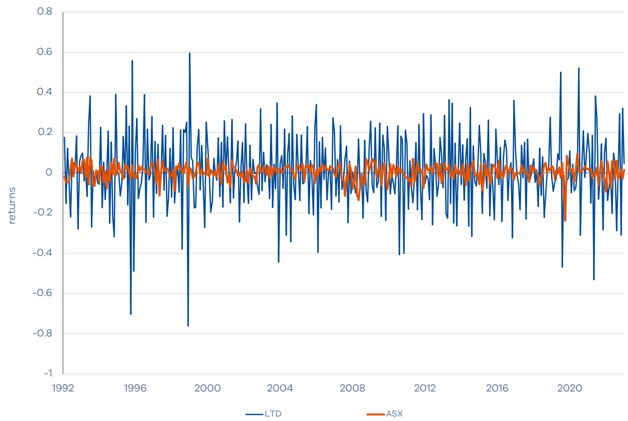
1 The views expressed in this paper are those of the authors and do not necessarily reflect the views of the Department of Treasury and Finance (DTF).

Figure 1: Land transfer duty

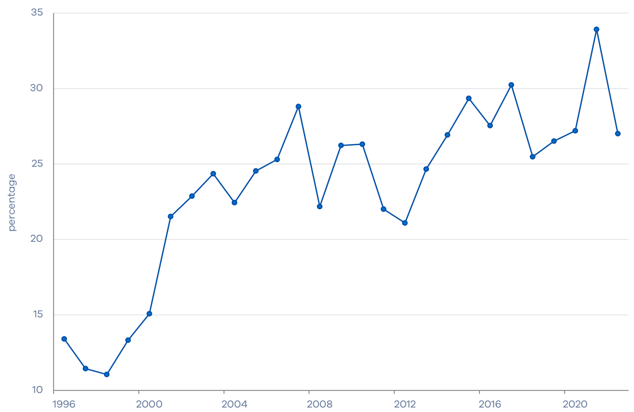
(a) FY 2023 Victorian tax revenue



(b) LTD return volatility



(c) LTD share of Victorian tax revenue



The forecasting of land transfer duty (LTD) in Victoria is challenging given this volatility and the cyclicality of the property market. The task becomes even more complex when considering broader economic factors that influence property transactions, such as interest rates, economic growth, and policy changes. This paper aims to address these complexities by employing regime switching models, specifically Markov switching models. This model class has gained prominence in recent literature given its ability to capture non-linearities and structural breaks.

Accurate revenue forecasts are critical for the Victorian Government's expenditure and budget planning, especially for significant revenue sources like LTD. The risks of forecast inaccuracies are amplified during economic disruptions. For instance, the COVID-19 pandemic severely impacted the Victorian property market, resulting in a 58 per cent decline in LTD revenue over the first six months of 2020. These factors underscore the potential value of models that can capture market shifts through regime switching.

Traditional linear models, such as vector auto-regression (VAR) (Sims, 1980), assume a stable relationship between variables over time. This may be unrealistic in a property market influenced by various factors which may cause sudden shifts in transaction volumes and prices. Markov switching models (Hamilton, 1989) have been used to model and forecast time series with distinct regimes, such as high and low-volatility periods. Interactions between property prices, sales volumes and LTD may differ during expansionary and contractionary periods, so we aim to capture any shifts and improve LTD forecasts.

Cointegration is vital for analysing multiple time series that exhibit long-run equilibrium relationships. The vector error correction model (VECM) extends the VAR to incorporate cointegration between non-stationary series, while also adjusting for short-term deviations from equilibrium. When combined with regime switching, VECMs have been shown to capture long-run relationships and shifts across states, providing superior forecasts in volatile markets (Lim and Tsiaplias, 2019, Ang and Bekaert, 2002, Krolzig, 2001) and across business cycles (Filardo, 1994).

Our results support the use of the MS-VECM when modelling and forecasting LTD. Our models identify two states: a contractionary short-lived state characterised by low LTD growth and heightened volatility; and an expansionary (normal) state with stronger LTD growth and lower volatility. In real time, the models identify the high volatility state associated with the onset of COVID-19. The MS-VECM consistently improves LTD forecasts across multiple horizons and loss functions relative to benchmarks.

We proceed as follows: Section 2 summarises the relevant literature on non-linear models. Section 3 details the models and their estimation. Section 4 presents estimation results and out-of-sample forecast performance, including an analysis over the COVID-19 period. Section 5 concludes the paper.

2. Literature review

The goal of this paper is to combine the Markov switching framework with a vector error correction model (VECM). We therefore seek a model that can jointly capture regime switches and long-run equilibrium dynamics.

Standard time series models assume a stable relationship between variables over time. This is despite tests which often suggest regime switches in macroeconomic and financial variables (Bai and Perron, 2003). This is important because unmodelled breaks can lead to forecast failure (Clements and Hendry, 1999).

Markov switching models (Hamilton, 1989) are a popular way to model structural breaks.[2] The regime is determined by an unobserved latent variable that follows a discrete time Markov Chain. Each regime has its own set of parameters, so the model is well-suited for capturing non-linear dynamics and regime shifts. Furthermore, the construction of a Markov Chain allows for the study of regime movements through time and the forecasting of future states.[3]

Markov switching models have been used to model expansions and contractions in business cycles, and shifts between bull and bear markets (Hamilton, 1994 and Guidolin and Timmermann, 2005). This is important, as macroeconomic shocks such as the COVID-19 pandemic can cause structural breaks, particularly in financial and property markets (Trefz, 2023 and Issam et al., 2024). Shen (2014) uses a Markov switching vector error correction model (MS-VECM) to examine housing market interactions with the stock market and consumer spending across different regimes. During economic expansions, rising housing prices positively affect consumer spending through the 'wealth effect'. However, in downturns, the housing market's influence on consumption can weaken or even become negative, exacerbating consumer pessimism and reducing spending.

When forecasting, Markov switching models have outperformed linear (state invariant) models (Ang and Bekaert, 2002, Filardo, 1994, and Krolzig, 2001). Markov switching models however do not always provide reliable forecasts as minor errors in predicting the regime can severely degrade performance, resulting in higher forecast errors than a random walk (Dacco and Satchell, 1999). Furthermore, determining the number of regimes is essential, but plagued by unidentified nuisance parameters under the null hypothesis of no structural change. This model class may therefore perform worse than state-invariant alternatives due to the large number of parameters, and the possibility of false break identification (Guidolin, 2009 and Guidolin et al., 2014).[4]

While Markov switching models offer advantages in capturing non-linearities and structural breaks, their forecasting performance relies heavily on accurate regime identification and lag selection. We integrate the Markov switching framework with a VECM to assess whether these models can outperform traditional state-invariant approaches in forecasting LTD.

3. Data and methodology

We perform two sets of analysis to identify regime switching ability and forecast accuracy. The first estimates MS-VECM models using the entire data set. We use this analysis to gain insights into the ability of the models to infer regime switches over the entire sample. In the second part of our analysis, we examine the out-of-sample forecast performance of our model. Section 3.1 introduces the data and performs tests for cointegration. Section 3.2 discusses the MS-VECM model and its estimation. Section 3.3 then describes the model selection criteria and out-of-sample forecast methodology. The results follow in section 4.

3.1 Data and preliminary analysis

This section describes the data, focusing on the variables that form the core of the forecasting models.

LTD is levied based on property value, making it a function of both house prices and the volume of property transfers. Accordingly, we define the following three-variable system:

$$x_t = [LTD_t \quad HVI_t \quad VOL_t]$$

where: LTD_t represents land transfer duty at time t, HVI_t the CoreLogic Home Value Index for Victoria, and VOL_t the volume of property transfers in Victoria.[5] Each variable is monthly, seasonally adjusted, and transformed into logarithmic form.

2 Other regime switching models include threshold models with discrete or smooth transitions between regimes (Tsay, 1998 and Chan and Tong, 1985). This model class was also examined, however threshold switching models performed poorly.

3 Krolzig (1997b) provides an excellent theoretical foundation for understanding and estimating Markov switching models.

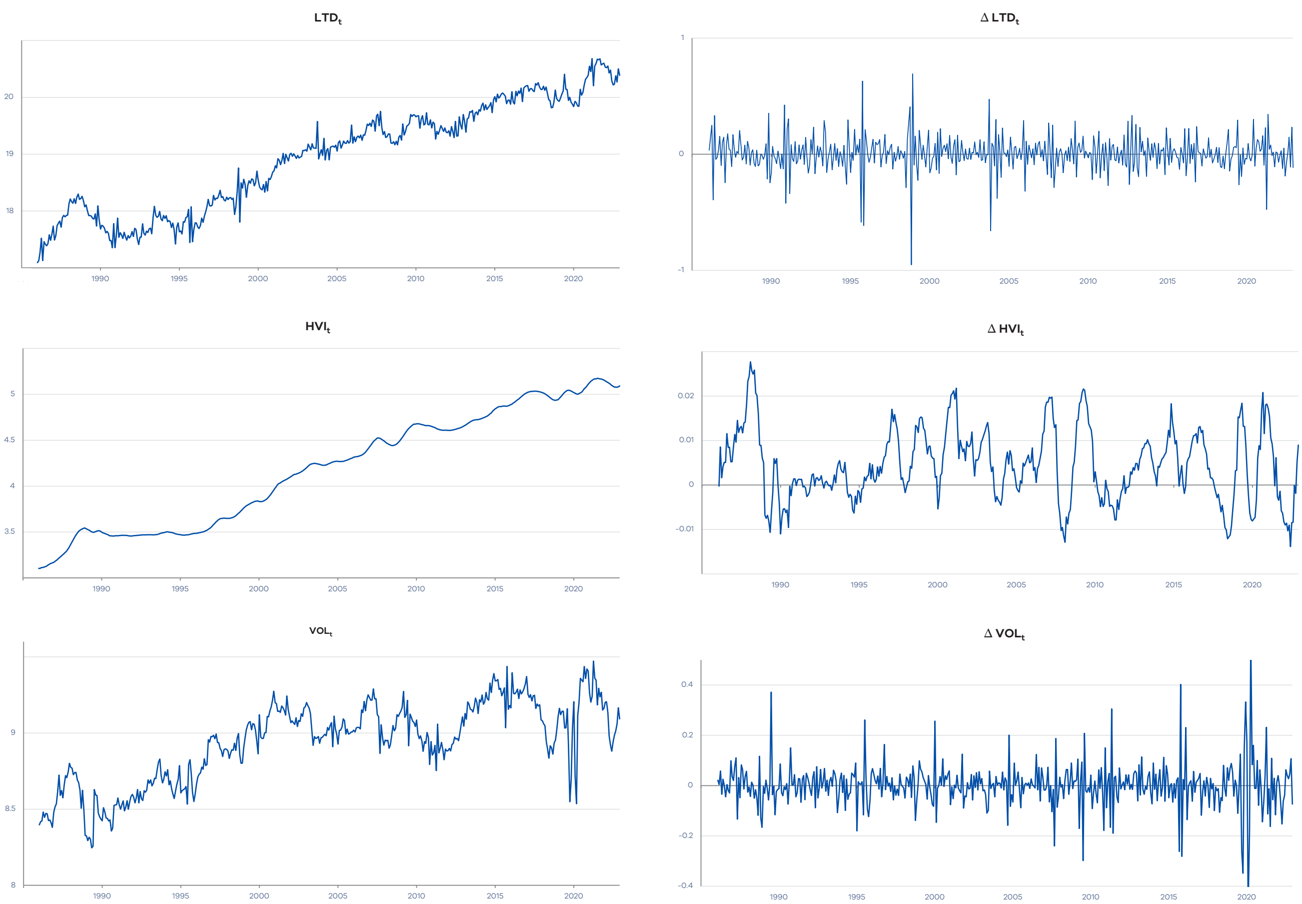
4 Recent advances in Markov switching estimation, such as Bayesian methods by Kim and Nelson (1999), have improved model robustness under uncertain economic conditions. Bayesian models are computationally demanding and are the subject of future research.

5 The CoreLogic Home Value Index (HVI) tracks changes in residential property values across regions in Australia. It is derived from a hedonic regression model that adjusts for property characteristics to accurately reflect market trends (CoreLogic Australia, 2023).

The data set spans July 1986 to June 2023, providing 444 observations. Figure 2 presents time series plots of x_t (in levels) and its first differences $\Delta x_t = x_t - x_{t-1}$. The plots reveal that volume was heavily impacted by COVID-19, with property sales plummeting 61 per cent in March and April 2020. This had significant effects on the Victorian housing market and may be evidence of a structural break.

Table 1 summarises the key descriptive statistics along with Augmented Dickey-Fuller (ADF) and ARCH tests (Engle, 1982). Results show that each series is $\mathrm{I}(1)$ in levels but $\mathrm{I}(0)$ once differenced. ARCH tests indicate time-varying volatility in LTD and VOL. While ARCH/GARCH errors (Bollerslev, 1986) are not directly modelled here, our Markov switching models (below) allow for state-dependent covariance matrices, capturing heteroskedasticity across regimes.[6]

Figure 2: Time series plots of x_t and Δx_t



6 This can occur because the model covariance is a weighted mixture, where the weights (regime probabilities) evolve through time.

Table 1: Descriptive statistics

SERIES	μ	σ	SKEW	KURT	ADF	ARCH
LTD_t	18.9439	0.9524	-0.1100	1.7229	0.9202	119.8286**
HVI_t	4.1898	0.6266	-0.0323	1.5675	11.3262	313.9819**
VOL_t	8.9350	0.2665	-0.4652	2.3793	0.2596	91.1425**
ΔLTD_t	0.0074	0.1536	-0.5762	9.0680	-36.9412**	59.9471**
ΔHVI_t	0.0045	0.0080	0.2775	2.7385	-3.501**	7.204
ΔVOL_t	0.0016	0.0887	0.4784	11.5881	-26.4065**	118.2388**

Note: All variables are in natural logs. The final two columns report the ADF and ARCH test statistics, with critical values of -1.941 and 7.815, respectively. Notations * and ** denote rejection of the null hypothesis at the 5 per cent and 1 per cent levels. ARCH tests are based on three lags of fitted AR(1) residuals.

Table 2 reports the trace test for cointegration between the $\boldsymbol{x}_t$ variables.[7] The rejection of the first null hypothesis ($r = 0$) and failure to reject the second ($r \leq 1$) suggests one cointegrating vector. The maximum eigenvalue test provides similar results. This supports one, state-invariant cointegrating vector in the regime switching models that follow. Despite this, we also consider a VAR in levels because it may outperform a VECM when the number of cointegrating vectors or their form is mis-specified (Clark, 2000).[8]

Table 2: Trace test

	EIGENVALUE	TEST STAT	CRITICAL VALUE	P-VALUE
$r = 0$	0.0532	33.8647	29.7976	0.0162
$r \leq 1$	0.0195	9.8816	15.4948	0.3287
$r \leq 2$	0.0028	1.2393	3.8415	0.3532

3.2 MS-VECM and smoothed probabilities

To develop the Markov switching VECM (MS-VECM), we commence with the state-invariant VECM. The VECM(p) is represented as follows:

$$\Delta x_t = \nu + \alpha\beta^T x_{t-1} + \sum_{k=1}^{p} \Gamma_k \Delta x_{t-k} + \varepsilon_t$$

where $\boldsymbol{\Pi} = \boldsymbol{\alpha\beta}^T$ is the error correction term split into two 3×r matrices: the adjustment term $\boldsymbol{\alpha}$ and the cointegrating vector(s) $\boldsymbol{\beta}$. The cointegrating vector contains r cointegration relations between the variables in $\boldsymbol{x}_t$, where r is the rank of $\boldsymbol{\Pi}$. $\boldsymbol{\nu}$ is a 3×1 vector of intercepts, $\boldsymbol{\Gamma}k$ are p autoregressive parameter matrices of dimension 3×3, and $\boldsymbol{\epsilon}t$ is a 3×1 vector of innovations at time t.

Markov switching vector error correction models (MS-VECMs) are an extension of traditional VECM models that allow for state-dependent parameters. The central feature of the MS-VECM is a latent, unobservable variable that governs which regime the model is in at any given time. This latent variable follows a Markov process, which dictates the transition probabilities between different regimes. Each regime is characterised by its own set of parameters, such as the speed of adjustment towards the long-term equilibrium, short-term dynamics, and error variance.

7 This form of Johansen test (Johansen, 1988) provides robust results, yet it has limitations. It can have low power in small samples (Cheung and Lai, 1993), though the long dataset (1986–2023) helps mitigate this issue. Additionally, sensitivity to lag length selection may affect the test's outcomes, so lag selection is optimised using Akaike and Schwarz information criteria (AIC and SIC) to enhance robustness.

8 The inverse roots from the VAR(4) model in levels are also within the unit circle. One root is close to but less than unity, which is supportive of cointegration, but also indicates that the VAR in levels will generate non-explosive forecasts.

The unobserved latent variable, S_t, takes on a finite number of values, representing each different regime. Given our modest sample size and the large number of parameters, our MS-VECM models only allow for a maximum of two regimes. The transition between these regimes occurs discretely at every step in time (in this case, every month). The switch is governed by a discrete time Markov chain – a stochastic process which specifies the probabilities of moving from one regime to another. The transition matrix of a two-state Markov chain, denoted $\boldsymbol{P}$, contains these probabilities where each element p_{ij} represents the probability of transitioning from regime i to j in the next period.

$$P = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix}$$

Markov chains are memoryless, meaning the probability of moving to a different state depends only on the current regime and not on any past states:

$$\Pr(s_{t+1} = j | s_t = i, s_{t-1}, s_{t-2}, \ldots) = \Pr(s_{t+1} = j | s_t = i) = p_{ij}$$

The two-state MS-VECM(p) with state dependent constant, adjustment speed to cointegrating equilibrium, heteroskedasticity, and short-run dynamics can be expressed as follows:

$$\Delta x_t = \nu_{s_t} + \alpha_{s_t} \beta^T x_{t-1} + \sum_{k=1}^{p} \Gamma_{k,s_t} \Delta x_{t-k} + \varepsilon_{s_t}$$

with state dependent parameters:

$$\alpha_{st} = \begin{cases} \alpha_1 \text{ if } s_t = 1 \\ \alpha_2 \text{ if } s_t = 2 \end{cases} \quad v_{st} = \begin{cases} v_1 \text{ if } s_t = 1 \\ v_2 \text{ if } s_t = 2 \end{cases}$$

$$\Gamma_{st} = \begin{cases} \Gamma_{k,1} \text{ if } s_t = 1 \\ \Gamma_{k,2} \text{ if } s_t = 2 \end{cases} \quad \varepsilon_{st} \sim \begin{cases} NID(\mathbf{0}, \mathbf{\Sigma}_1) \text{ if } s_t = 1 \\ NID(\mathbf{0}, \mathbf{\Sigma}_2) \text{ if } s_t = 2 \end{cases}$$

An alternative form of regime switching examined in this paper is the Markov switching autoregressive invariant vector error correction model (MSAI-VECM). This model constrains the short-run dynamics (autoregressive terms) to remain constant across regimes, unlike the full MS-VECM, where all parameters may vary by regime. The MSAI-VECM is therefore a restricted, nested version of the MS-VECM. It is primarily implemented to reduce the number of estimable parameters, thereby improving the likelihood that the EM algorithm converges. The MSAI-VECM is written as follows:

$$\Delta \boldsymbol{x}_t = \mathbf{v}_{s_t} + \boldsymbol{\alpha}_{s_t} \boldsymbol{\beta}^T \boldsymbol{x}_{t-1} + \sum_{k=1}^{p} \boldsymbol{\Gamma}_k \, \Delta \boldsymbol{x}_{t-k} + \boldsymbol{\epsilon}_{s_t}$$

We estimate our models using the following two-step procedure:

1. Estimate the number of cointegrating vectors and their coefficients using the Johansen procedure. The cointegrating vector $\boldsymbol{\beta}$ is therefore assumed to be constant across regimes, implying a state-invariant equilibrium relationship between LTD_t, HVI_t, and VOL_t.
2. Estimate the remaining parameters of the MS-VECM (including the regime transition matrix P) via maximum likelihood (MLE). We use the Expectation Maximisation (EM) algorithm. Bayesian methods could also be used for estimation as described by Kim and Nelson (1999).[9]

The EM algorithm estimates both the VECM parameters and regime probabilities via a smoothing algorithm. Once the parameters of the MS-VECM are estimated, the smoothed probabilities of being in each regime can be generated using the Hamilton filter or similar smoothing algorithms. These probabilities represent the likelihood that the system was in a particular regime at each point in time, given the entire dataset. The smoothed probabilities, denoted as $P\,(st = j|x_t)$, take into account both past and future observations to provide a more accurate estimate of regime transitions. This paper employs the reverse recursion algorithm of Kim (1994) to smooth filtered state probabilities.

Testing for regime switching in the MS-VECM framework is complicvated by the presence of unidentified nuisance parameters (the elements of P) under the null hypothesis of no regime switching. We therefore follow Psaradakis and Spagnolo (2003), who estimate the linear (state-invariant) VECM and the MS-VECM. If the MS-VECM has a lower AIC, this provides evidence in support of regime switching.

3.3 *Model selection via cross-validation and out-of-sample forecasting*

When estimating the MS-VECM over the full sample (July 1986 to June 2023), we use the SIC to determine the optimal number of lags. However for out-of-sample forecasting, information criteria may be sub-optimal (Krolzig, 1997a). We therefore follow Sarno et al (2004) and use cross-validation. We select the number of lags that minimise the forecast error over a training/cross-validation sample. We focus on the mean absolute percentage error (MAPE) for 12-month ahead cumulative forecast errors. However, mean squared errors (MSE) are also included for robustness.

9 The conditional log-likelihood function for a two regime MS-VECM is given by:

$$L(\Delta \mathbf{x}_t \mid \Delta \mathbf{x}_{t-1}, \xi_{s_t}, s_t) \propto \prod_{t=1}^{T} \left\{ \sum_{s_t=1}^{2} \left((2\pi)^{-1/2} |\mathbf{\Sigma}_{s_t}|^{-1/2} \right) \exp\left(-\frac{1}{2} \mathrm{tr} \left[\mathrm{vec}(\epsilon_{s_t})^{T} (\mathbf{\Sigma}_{s_t} \otimes \mathbf{I}) \mathrm{vec}(\epsilon_{s_t}) \right] \right) \right\} \Pr(\xi_{s_t})$$

where $\xi_{s_t} = \{\mathbf{v}_{s_t}, \boldsymbol{\alpha}_{s_t}, \boldsymbol{\Gamma}_{s_t}, \mathbf{\Sigma}_{s_t}, p_{IJ}\}$ contains all information about the realisation of the Markov chain, and $\xi t, {}_{s_t}$ are the conditional regime probabilities calculated by the recursive filtering and smoothing algorithms discussed in Krolzig (1997b). *tr*, *vec*, and $\otimes$ are the trace operator, vectorisation operator, and Kronecker product, respectively. The model is therefore a mixture of n distributions, where n = 2 is the number of states.

The general h-step ahead cumulative forecast loss is:

$$e(x_{t+h}) = \sum_{k=1}^{h} \widehat{x_{t+k}} - x_{t+k}$$

and the h-step ahead cumulative MAPE over the forecasting period [τ, T] is:

$$MAPE(x_{t+h}, \tau, T) = \frac{1}{T - \tau + 1} \sum_{t=\tau}^{T} \left| \frac{e(x_{t+h})}{\sum_{k=1}^{h} x_{t+k}} \right|$$

The h-step ahead cumulative MSE over the forecasting period [τ, T] is:

$$MSE(x_{t+h}, \tau, T) = \frac{1}{T - \tau + 1} \sum_{t=\tau}^{T} e(x_{t+h})^2$$

The cross-validation undertakes a forecasting exercise through the training sample by estimating a model using data from time 1 to time τ, and calculating the conditional h-month ahead forecast $\hat{x}_{\tau+h}$. Then the window steps forward a sample (i.e. estimating from 2 to τ + 1) and forecasts $\hat{x}_{(\tau+1)+h}$. This procedure iterates until the final forecast $\hat{x}_T$ has been computed. The cross-validation repeats this exercise for every lag specification and chooses the model which minimises the h-step ahead cumulative MAPE:

$$argmin_p \; MAPE(x_{t+h}, \tau, T)$$

where forecasts $\hat{x}_{t+h}$ are computed using a linear or MS model with lag length p.

Section 4 selects models that minimise the 12-step-ahead cumulative MAPE over the training sample (h = 12). These optimal specifications are referred to as the CV(12) models. For robustness, the specifications and forecast performance of CV(3) models, which minimise the 3-step-ahead cumulative MAPE (h = 3), are provided in Appendix A.

Our initial estimation employs a 25-year window from July 1986 to June 2011. The six years from July 2011 to June 2017 serve as the training sample. After selecting the number of lags, the model specifications are fixed. We then used the rest of the sample (July 2017 to June 2023) to assess the out-of-sample forecasts. Like the cross-validation exercise, the 25-year estimation window is rolled forward each period. The model is re-estimated and forecasts generated recursively over 1 to 12 month horizons conditional on the information set.

Forecast performance is evaluated using the Model Confidence Set (MCS) of Hansen et al (2011). This procedure is based on the cumulative loss differential $e(x_{t+h})$ and the MAPE associated with this loss:

$$d_{ij,t} = \frac{|e_i(x_{t+h})| - |e_j(x_{t+h})|}{\sum_{k=1}^{h} x_t}$$

where ei and ej are the forecast errors from the competing models i and j, respectively. The MCS procedure identifies a set of models that are statistically indistinguishable from the 'best' model at a given confidence level. The MCS test statistic is based on the forecast errors of each model in the set M under investigation. The optimal set $M^*_\alpha \subseteq M$ includes the best performing models at the chosen confidence level α. By computing the test statistic and p-value for each model relative to the 'best' model, models are iteratively removed from the set if their p-value falls below α.[10] If multiple models remain in M^*_α, it indicates failure to reject the null that they have equal predictive ability.

4. Results

This section is divided into four parts: Section 4.1 studies the full sample estimation results for the MS-VECM. We then consider forecast performance. This commences in section 4.2 which performs the cross-validation, and is followed by the out-of-sample (OOS) forecasts in section 4.3. Finally, section 4.4 considers a subset of the out-of-sample period, focusing on the COVID-19 period and the ability of the model to identify regime switches in real time.

4.1 Full sample estimation results

Here we compare estimation results from the VECM and MS-VECM. The linear VECM provides a useful baseline for comparison to the MS-VECM, which allows for state-dependent intercepts, speed of adjustment, auto-regressive lags and volatility. The estimated parameters for the two models are presented in Table 3.

10 The relevant t-statistics are:

$$t_{ij} = \frac{\overline{d_{ij}}}{\widehat{\sigma_{ij}}}, \quad \text{for } i, j \in M$$

where $\hat{\sigma}_{ij}$ is the estimated standard deviation of $\bar{d}_{ij}$. The MCS uses the range test statistic:

$$T_R = \max_{i,j \in M} |t_{ij}|$$

Table 3: Estimation results

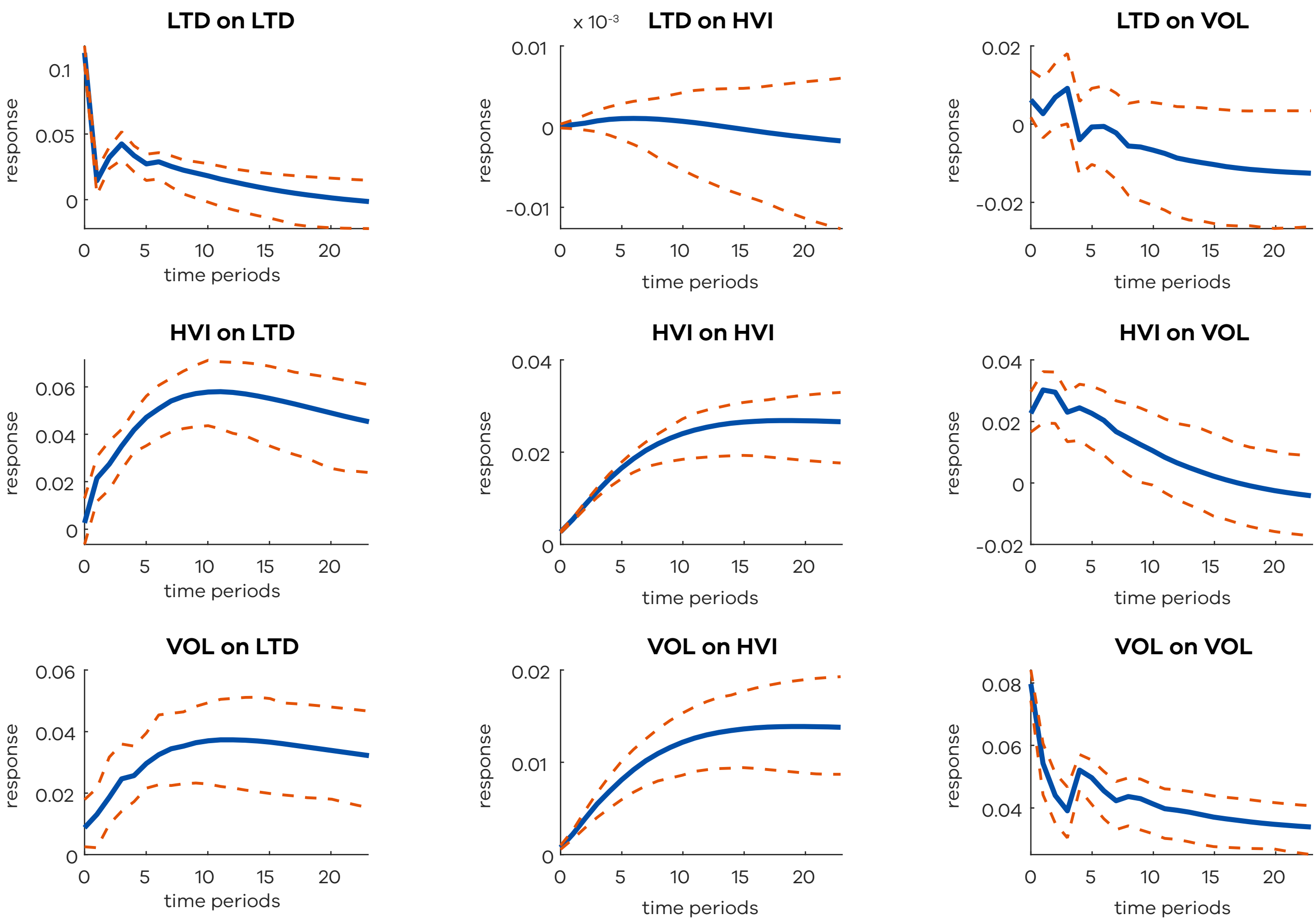
	VECM(3)			MS-VECM(1)		
	ΔLTD_t	ΔHVI_t	ΔVOL_t	ΔLTD_t	ΔHVI_t	ΔVOL_t
Regime dependent intercepts						
ν_1	2.269	0.003	0.697	1.775	0.008	0.374
	(4.93)	(0.30)	(2.13)	(6.37)	(1.07)	(2.58)
ν_2				5.270	0.063	-0.527
				(10.66)	(5.86)	(-1.21)
Regime dependent error correction						
$\alpha_1(\times 10^2)$	-2.644	-0.003	-0.802	-2.802	-0.012	-0.587
	(-4.95)	(-0.26)	(-2.11)	(-6.38)	(-1.02)	(-2.57)
$\alpha_2(\times 10^2)$				-8.331	-0.099	0.846
				(-10.69)	(-5.84)	(1.23)
β	7.932	-11.185	-1.909	7.104	-9.368	-3.566
Regime 1 short-run dynamics						
ΔLTD_{t-1}	-0.669	0.001	0.050	-0.346	0.002	-0.031
	(-12.26)	(0.54)	(1.28)	(-8.99)	(1.50)	(-1.56)
ΔHVI_{t-1}	6.721	0.903	5.791	3.174	0.953	0.212
	(3.44)	(18.86)	(4.17)	(5.74)	(63.76)	(0.74)
ΔVOL_{t-1}	0.032	0.009	-0.390	0.132	0.007	-0.170
	(0.48)	(5.50)	(-8.15)	(1.72)	(3.20)	(-4.24)
ΔLTD_{t-2}	-0.410	0.001	0.088			
	(-7.25)	(0.60)	(2.19)			
ΔHVI_{t-2}	0.754	0.190	-2.246			
	(0.29)	(3.02)	(-1.23)			
ΔVOL_{t-2}	0.026	0.006	-0.350			
	(0.36)	(3.18)	(-6.90)			
ΔLTD_{t-3}	-0.118	0.001	0.122			
	(-2.57)	(1.22)	(3.75)			
ΔHVI_{t-3}	-1.010	-0.198	-4.283			
	(-0.52)	(-4.15)	(-3.10)			
ΔVOL_{t-3}	0.032	0.003	-0.282			
	(0.47)	(1.55)	(-5.79)			
Regime 2 short-run dynamics						
ΔLTD_{t-1}				-0.366	0.004	-0.046
				(-7.87)	(3.83)	(-1.13)
ΔHVI_{t-1}				5.116	0.828	-2.062
				(5.02)	(37.20)	(-2.30)
ΔVOL_{t-1}				-0.204	0.005	-0.232
				(-3.37)	(3.58)	(-4.35)

	VECM(3)			MS-VECM(1)		
	ΔLTD_t	ΔHVI_t	ΔVOL_t	ΔLTD_t	ΔHVI_t	ΔVOL_t
Regime 1 covariance (×10²)						
ΔLTD_t	1.255	0.001	0.070	0.822	0.003	0.068
ΔHVI_t	0.001	0.001	0.006	0.003	0.001	0.004
ΔVOL_t	0.070	0.006	0.635	0.068	0.004	0.222
Regime 2 covariance (×10²)						
ΔLTD_t				2.915	-0.011	0.221
ΔHVI_t				-0.011	0.001	0.009
ΔVOL_t				0.221	0.009	2.268

Note: Full sample model estimates. Lag selection is based on SIC. Markov switching model allows for state dependent constants, autoregressive terms, adjustment speed to equilibrium, and heteroskedasticity. Table 4 reports the diagnostics and regime dynamics.

We commence with an analysis of the VECM, where the speed of adjustment α_1 measures how the variables x_t adjust to deviations from the long-run equilibrium. Our model estimates imply a long-run LTD growth rate of 0.673 per cent every month. Generalised impulse response functions (GIRF) with 95 per cent confidence intervals are plotted in Figure 3.

Figure 3: VECM(3) GIRFs



A shock to LTD causes a brief positive impact on LTD revenues, which quickly decays to zero. In contrast, shocks to HVI and VOL show much more persistent effects, with price and volume shocks increasing LTD revenues by around 6 per cent and 4 per cent, respectively, lasting for over 24 months. This suggests that price shocks have a more substantial impact on LTD than volume shocks, aligning with the progressive tax structure. LTD shocks have no significant effect on property prices and transaction volumes. A shock to HVI also has a positive short-term effect on VOL, potentially indicating a strong market that encourages speculative activity, and thus increased volume. Similarly, a shock to VOL leads to a significant and lasting increase in HVI, suggesting volume shocks may signal rising demand that pushes prices higher.

Model diagnostics in Table 4 demonstrate the MS-VECM's substantially lower AIC, which provides evidence in favour of regime switching.[11] The MS-VECM identifies a persistent, low-volatility state with steady state LTD growth of 1.078 per cent every month (state one) and a less persistent, high-volatility state with slower LTD growth of 0.451 per cent every month (state two). The volatility of LTD and VOL in regime two is higher than regime one by over 4 and 10 times respectively. The MS-VECM estimate of α_1 aligns closely with the VECM estimate. However, α_2 suggests a faster return to equilibrium in the contractionary state.

Table 4: Model diagnostics

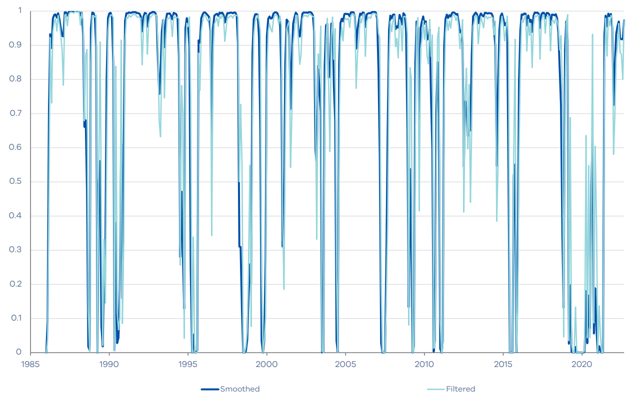
	VECM(3)	MS-VECM(1)
SIC	-5417.86	-5602.24
AIC	-5564.98	-5724.98
LL	2818.49	2892.49
ADF	-21.1457**	-23.7975**
ΔLTD_1	0.673%	1.078%
ΔLTD_2		0.451%
p_{11}		0.9183
p_{22}		0.7429

Note: Model lag and threshold variable selection chosen by SIC. Markov switching model allows for state dependent constants, autoregressive terms, adjustment speed to equilibrium, and heteroskedasticity. The ADF tests are run on model residuals. Indicators * and ** represent the rejection of the null hypothesis at the 5 per cent and 1 per cent levels, respectively. ΔLTD_k represents the long-run LTD growth rate of the model in state k.

The transition matrix suggests that state one – the expansion state – is more persistent than state two. The MS-VECM is expected to enter a contractionary phase every 12.24 months, with each contraction lasting 3.89 months on average. The ergodic (steady state) probabilities, which represent the expected long-run frequency of being in each regime, are calculated as: $\boldsymbol{\pi P} = \boldsymbol{\pi}$, $\boldsymbol{\pi} = [\pi_1, \pi_2] = [0.759, 0.241]$. Thus, the housing market is expected to be in the first state 75.9 per cent of the time.

The smoothed and filtered probability plot for regime one over the last 20 years of the sample is shown in Figure 4. Switches into the low-growth high-volatility state (state 2) are evident during the Australian recession in the early 1990s, the global financial crisis in 2008, and the COVID-19 pandemic in 2020. It is worth noting that during COVID-19, the model remains in state 2 for an unprecedented 9-month period, nearly three times longer than expected from the transition matrix.

Figure 4: State 1 probabilities

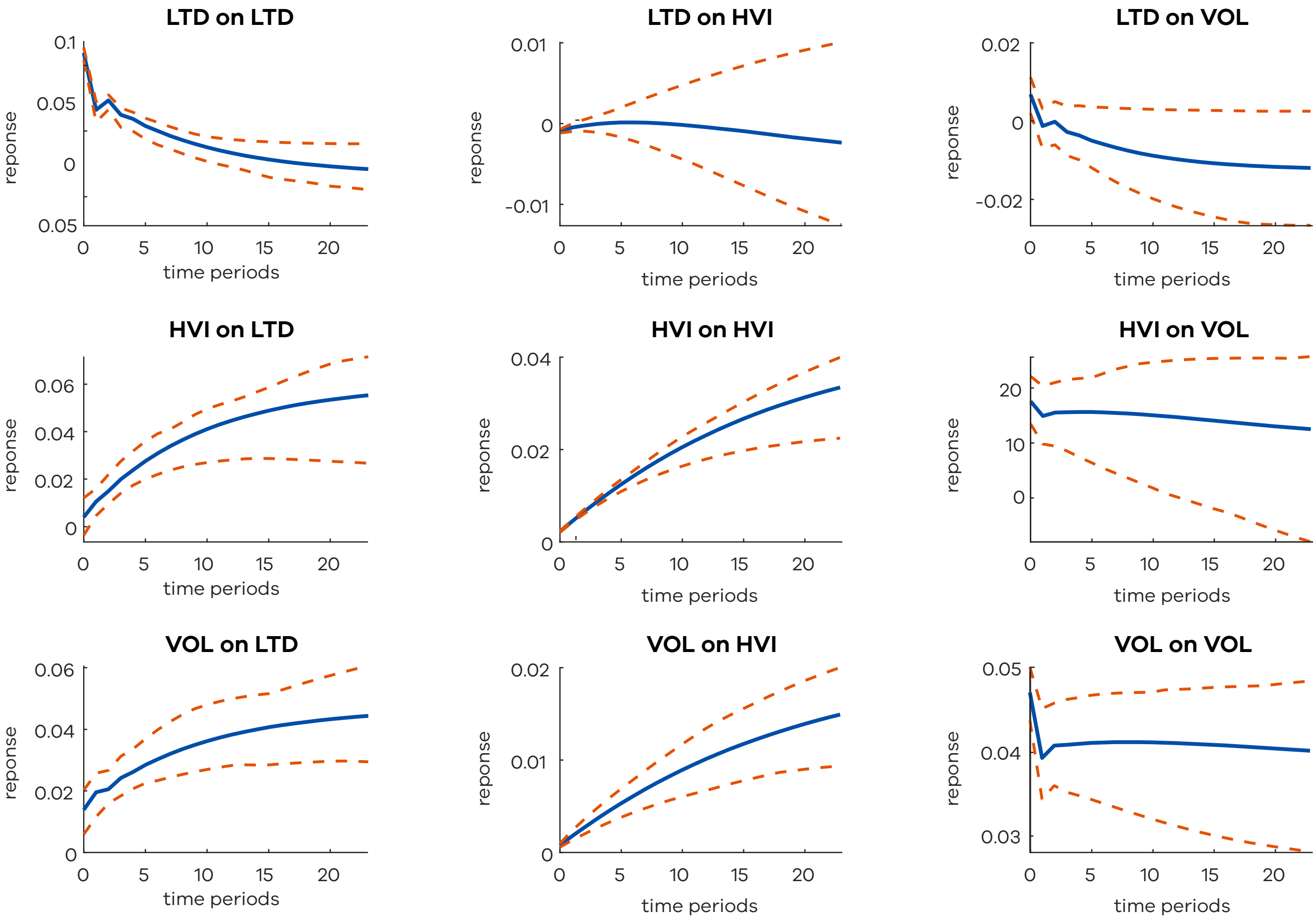


Finally, Figures 5a and 5b display the 24-step-ahead GIRFs and 95 per cent confidence bands for the two MS-VECM states. In regime one, price and volume shocks have persistent impacts on LTD revenues, similar to the VECM. Prices respond positively to volume shocks, indicating a strong market during expansionary periods, which encourages speculation and increases volume. The short-run dynamics shift in regime two, where LTD reacts more quickly to volume and price shocks, though with similar magnitude. Notably, an LTD shock causes a significant price decline and volume increase, suggesting there may be greater sensitivity of the housing market to policy changes during contractions.

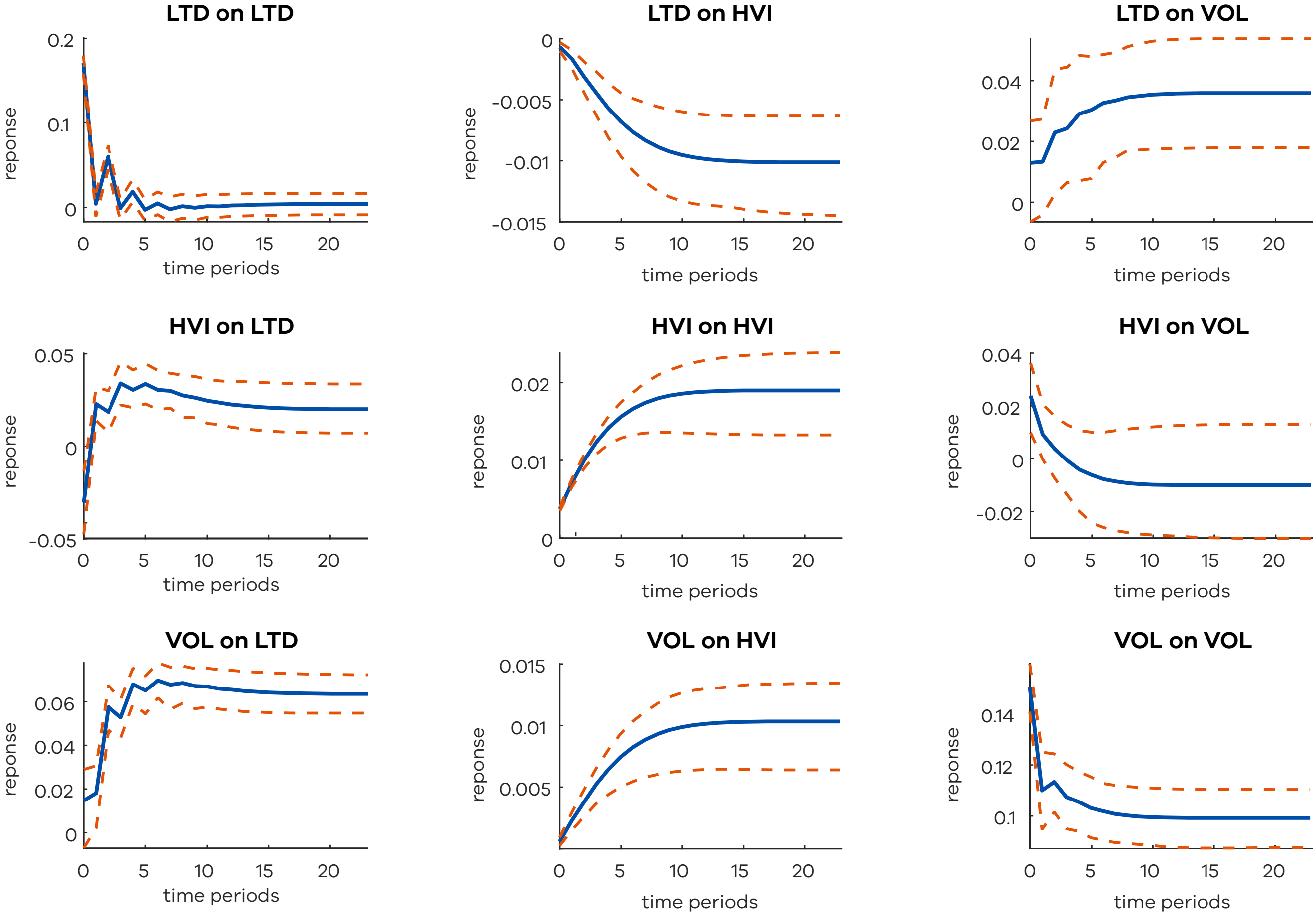
11 Note the Markov switching model selects one autoregressive lag, compared to the linear VECM's three lags. This is likely due to the SIC penalising the additional parameters in the state dependent models.

Figure 5: MS-VECM(1) GIRFs

(a) State 1: high growth, low volatility



(b) State 2: low growth, high volatility



In summary, the MS-VECM provides a plausible characterisation of the dynamics, identifying switches between a high-LTD-growth low-volatility state and a low LTD-growth high-volatility state. The next three sections will examine whether allowing for regime switches improves LTD forecasts.

4.2 *Cross-validation*

Our primary goal is to produce accurate and reliable forecasts of LTD. To evaluate forecast performance of the MS-VECM model, we perform a comprehensive 'horse race' between the following models: VECM, VAR (in levels), dVAR (VAR in first differences), dAR (univariate AR in first differences), MS-VECM, MSAI-VECM, MS-VAR (in levels), MS-dVAR (first differences), and MS-dAR (univariate).

All models require lag length selection and, as discussed above, traditional criteria such as SIC may not yield optimal forecasts over longer horizons. To address this, the cross-validation (CV) approach in Section 3.3 is implemented. We use a 25-year estimation window, a 6-year training period (July 2011 to June 2017) and the last 6-year period (July 2017 to June 2023) for out-of-sample (OOS) forecasting. Experimentation revealed that a 25-year estimation window is essential for the MS models to identify regime switches. We are therefore limited to small sample sizes for the CV and OOS periods, which may affect our ability to statistically detect differences between models.[12] Our CV analysis optimises lag length by considering lags of 1 to 10 for the linear models and 1 to 4 for the MS models (as the additional regime increases each models' dimensionality).

The optimal model specifications selected via CV are outlined in Table 5. Cross-validation typically favours fewer lags for models in levels, while selecting more lags for those in first differences, as demonstrated by the linear VAR and dVAR results. This pattern contrasts with AIC/SIC, which tends to select the same number of lags across model types and restricts lag length to a maximum of four. The benchmark model used for performance comparison is the VECM(4), commonly adopted by the Department of Treasury and Finance (DTF) due to its consistent out-of-sample accuracy.[13]

Table 5: Cross validation results

		STATE DEPENDENT:			
SPECIFICATION	**LAGS**	ν	α	Γ_k	Σ
Selected models					
VECM	4				
VAR	2				
dVAR	9				
dAR	12				
MS-VECM	3	Yes	Yes	Yes	Yes
MSAI-VECM	4	Yes	Yes	No	Yes
MS-VAR	2	Yes	Yes	Yes	Yes
MS-dVAR	4	Yes	Yes	Yes	Yes
MS-dAR	6	Yes	Yes	Yes	Yes
Model combinations					
AVG-3 (10%)					
AVG-12 (10%)					
AVG-3 (25%)					
AVG-12 (25%)					
MS-Combo					

Note: Forecasting models specification determined by cross-validation approach, choosing the number of lags which minimise a model's 12-month ahead cumulative MAPE over the training sample (7/2011-6/2017). Linear and MS models consider up to 12 and 6 lags, respectively. AVG-n (10 per cent) and AVG-n (25 per cent) models refer to model averages for the n-horizon $M^*_{0.10}$ and $M^*_{0.25}$, respectively, determined during the training sample (7/2011-6/2017). MS-Combo is the average of the best performing regime switching models during the training sample (7/2011-6/2017) determined by the MCS. See Appendix A for a list of models in each combination.

12 MS models without sufficient data suffer from two main issues: (i) the EM algorithm required for MLE parameter estimation is sensitive to initial values and may converge to local maxima rather than the global maximum, resulting in incorrect parameter estimates; (ii) the MS models may fail to detect structural breaks, causing the EM algorithm to fail to converge to well-behaved estimates. Window sizes smaller than 25 years commonly encountered these issues.

13 While the CV(12) process selects the VECM(8) as the optimal model over the training sample, the VECM(4) achieves superior out-of-sample performance. Forecast tables with results for the VECM(8) are available upon request.

Our CV exercise also considers model combination forecasts. The AVG-n (10 per cent) and AVG-n (25 per cent) models are simple average forecast combinations, including models from the 10 per cent and 25 per cent MCS for the respective n-month horizon. Additionally, the MS-Combo model represents a simple average of the best performing regime switching models identified by the MCS during the training sample. Independent of the confidence level or forecast horizon, the same three models are consistently selected (MS-VECM, MSAI-VECM and MS-VAR). Further details are provided in Appendix A.

4.3 Forecast performance

Table 6 provides the cumulative MAPE and MSE for the selected models and model combinations determined via cross validation. The best model for each horizon is underlined. Forecast performance varies significantly, with 12-month MAPEs ranging from 18.72 per cent for the least accurate model (MS-dVAR) to 10.25 per cent for the best performing model (MS-VECM).

Table 6: Forecast evaluation (7/2017-6/2023)

	MAPE (%)				MSE (×10^{-16})			
HORIZON	3	6	9	12	3	6	9	12
VECM	7.34*	8.12*	10.15**	11.53*	29.45*	134.12**	456.09**	1 092.85**
VAR	7.36**	8.43**	9.92**	11.24**	31.45**	147.31**	433.31**	987.09**
dVAR	7.47*	8.43**	11.52	13.79	31.43	153.48*	571.5*	1 423.0*
dAR	10.9	12.71	15.15	16.52	54.94	299.48	996.84	22 99.15*
MS-VECM	7.11**	7.88**	9.16**	10.25**	27.61**	124.83**	**393.96****	**937.0****
MSAI-VECM	**6.77****	**7.33****	9.74**	11.3**	**25.23****	**113.26****	410.56**	1 021.85**
MS-VAR	7.4**	8.64**	10.32**	12.05	31.95**	158.69**	480.92**	1 092.69*
MS-dVAR	12.29	14.8	16.93	18.72	72.51	432.77	1 382.07	3 251.48*
MS-dAR	10.69	12.55	14.98	16.45	52.62	282.27	981.93*	23 67.28*
AVG-3 (10%)	6.93**	7.88**	9.72**	11.03**	27.7**	129.16*	425.06**	10 07.78**
AVG-12 (10%)	7.37**	8.23**	10.18**	12.09**	28.45**	134.63**	456.78**	11 25.94**
AVG-3 (25%)	6.9**	7.85**	9.59**	10.83**	27.55**	128.03**	411.49**	969.1**
AVG-12 (25%)	6.96**	7.91**	9.52**	10.81**	27.59**	128.92**	407.33**	954.67**
MS-COMBO(12)	6.88**	7.76**	9.47**	10.74**	26.9**	125.45**	406.73**	963.59**

Note: The lowest loss for each horizon and loss function is underlined. **, * represent the model's inclusion in the MCS at 25 per cent $(M^*_{0.25})$ and 10 per cent $(M^*_{0.10})$levels respectively.

Note that $(M^*_{0.25}) \subseteq (M^*_{0.10})$. AVG-$n$ (10 per cent) and AVG-n (25 per cent) models refer to model averages for the n-horizon $(M^*_{0.10})$ and $(M^*_{0.25})$ determined during the training sample (7/2011-6/2017), respectively. MS-Combo is a simple average of the best performing Markov switching models. See Appendix A for a list of models in each combination. MSAI-VECM has state invariant autoregressive terms; all other parameters are state dependent.

The VECM(4), which serves as DTF's benchmark model, performs well out-of-sample, surpassing 6 of the 13 competing models at the critical 12-month horizon under the MAPE criterion. The linear VECM is included in the 10 per cent MCS $(M^*_{0.10})$ across all horizons and in the 25 per cent MCS $(M^*_{0.25})$ at the 9-month horizon under the MAPE loss, highlighting its robustness. Among the Markov switching models, the MS-VECM and MSAI-VECM achieve the lowest forecast errors across all horizons by both MAPE and MSE measures. Specifically, the MS-VECM outperforms the benchmark VECM by over 11 per cent, with statistically significant improvements in the 3, 6, and 12-month forecasts, as shown by its inclusion in $M^*_{0.25}$ using MAPE loss.

The MSAI-VECM and MS-VAR also outperform the VECM for the 3 and 6-month horizons, reflecting strong out-of-sample accuracy. The MS-dVAR and MS-dAR models forecast poorly, especially for shorter horizons. Despite MS-VECM and MSAI-VECM models achieving the lowest MSE, the inclusion of most models in the MCS at the 25 per cent level is likely due to the low power of the MSE loss function.

Finally, all model combinations are in $M^*_{0.25}$ under both loss functions, with none outperforming the MS-VECM or MSAI-VECM.

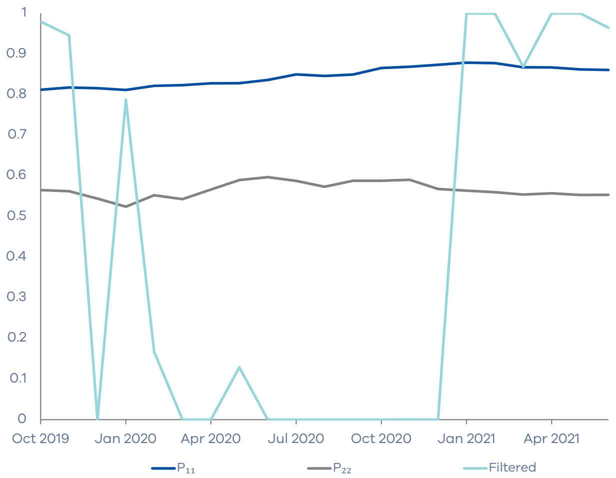
In summary, the Markov switching VECM models outperform a range of benchmarks including the linear VECM, across multiple horizons. A model combination consisting of Markov switching models also performs equally well.

4.4 COVID-19 analysis

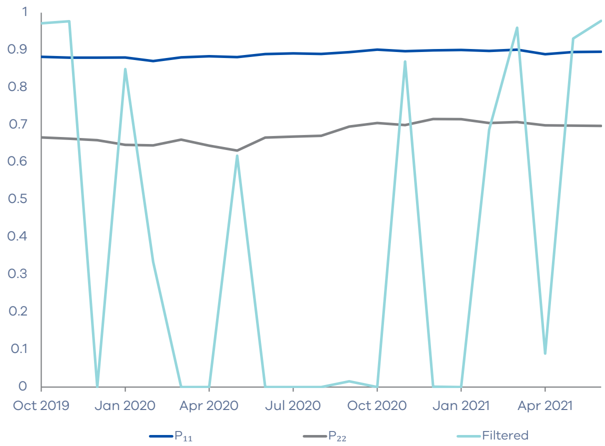
We now assess performance of regime-switching models during the pandemic (March 2020 to August 2021). First, we examine whether the MS models identify a structural break in real time. We then examine their out-of-sample forecast performance. Figure 6 plots filtered probabilities for the MS-VECM, MSAI-VECM and MS-VAR models.

Figure 6: MS model filtered regime 1 probabilities during COVID-19

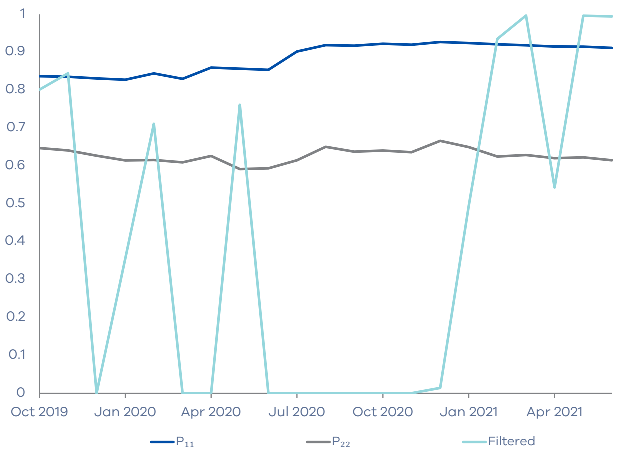
(a) MS-VECM(3)



(b) MSAI-VECM(4)



(c) MS-VAR(2)



The probabilities reflect the model's current assessment of which regime the system is in, conditional on the information at each point in time. To illustrate, the probability in March 2020, is the probability of being in state 1 from the model estimated using 25 years of data ending March 2020. The probability for April 2020, is based on the data set ending April 2020 etc. These probabilities therefore reflect a real-time assessment as the models roll through the estimation windows. The noisy probabilities throughout the sample suggest it's prudent to use a range of MS models when evaluating a regime switch in real time. Nonetheless, all model probabilities drop to zero in early 2020, indicating a switch to the low-growth high-volatility regime. This aligns with the smoothed probabilities from the full sample analysis shown in Figure 4.

Table 7 reports forecast errors for each model (selected via CV in section 4.2) from March 2020 to August 2021. The results are mixed with no model or combination clearly dominant across all horizons and loss functions. The MS-dVAR and MS-dAR models achieve the lowest 12-month forecast errors, with MAPEs of 13.31 per cent and 14.83 per cent, respectively, representing improvements of 23.68 per cent and 14.97 per cent over the linear VECM. The VECM though remains in $M^*_{0.25}$ across all horizons under both loss criteria, indicating that no model statistically outperforms the benchmark. Importantly, the MS-VECM has a lower MAPE and MSE than the VECM across all horizons, and it remains in the MCS at the 25 per cent level for horizons 6, 9 and 12.

Table 7: COVID-19 Forecast Evaluation (3/2020-8/2021)

	MAPE (%)				MSE (×10−16)			
HORIZON	**3**	**6**	**9**	**12**	**3**	**6**	**9**	**12**
VECM	9.08*	12.51**	16.85**	17.44**	28.29**	204.62**	801.52**	1 449.29**
VAR	8.58	11.56**	16.26**	18.64	28.33	200.19**	795.85**	1 589.17
dVAR	9.66	13.06	18.55	20.57	34.92	250.15	978.48	1 852.44
dAR	14.40	17.88	17.29*	15.14**	70.89	426.61	966.70*	1 376.60**
MS-VECM	8.31	11.32**	15.18**	15.46**	28.74	201.59**	773.57**	1 131.01**
MSAI-VECM	8.06**	11.84**	16.83*	17.76	25.65**	191.87**	786.51**	1 460.60*
MS-VAR	8.70**	11.89**	15.75**	17.69	27.36**	194.79**	773.17**	1 448.16
MS-dVAR	14.16	16.05**	16.78**	**13.31****	73.92	438.71*	1 153.97**	**1 128.83****
MS-dAR	14.32	18.03	17.35*	14.83**	69.38	433.27	971.48	1 313.87**
AVG-3 (10%)	7.90*	11.35	16.47*	17.75	26.49*	199.05	795.01*	1 438.71
AVG-12 (10%)	9.63	12.58**	**14.51****	15.19**	33.22*	235.15**	**757.39****	1 147.59**
AVG-3 (25%)	**7.69****	**11.09****	16.14*	17.40	**25.32****	**190.10****	766.34**	1 379.00
AVG-12 (25%)	7.85**	11.19**	15.95**	17.38	25.61**	190.34**	766.00**	1 377.28
MS-COMBO	7.72**	11.12**	15.9**	16.97	25.58**	190.65**	763.78**	1 319.80

Note: The lowest loss for each horizon and loss function is underlined. **, * represent the model's inclusion in the MCS at 25 per cent $(M^*_{0.25})$ and 10 per cent $(M^*_{0.10})$levels respectively.

Note that $(M^*_{0.25}) \subseteq (M^*_{0.10})$. AVG-$n$ (10 per cent) and AVG-n (25 per cent) models refer to model averages for the n-horizon $M^*_{0.10}$ and $M^*_{0.25}$ determined during the training sample (7/2011-6/2017), respectively. MS-Combo is a simple average of the best performing Markov switching models.
See Appendix A for a list of models in each combination. MSAI-VECM has state invariant autoregressive terms; all other parameters are state dependent.

The AVG-12 (10 per cent) and AVG-3 (25 per cent) are the best performing model combinations, achieving the smallest forecast errors at the 3, 6, and 9-month horizons. These findings highlight the effectiveness of model combinations during periods of economic turbulence. Simpler models in first differences, such as the MS-dVAR, MS-dAR and dAR, also perform relatively well, despite generally poor performance in the overall out-of-sample period (see section 4.3). This may be due to a range of factors including i) their simplicity, which may reduce the risk of over-fitting in highly volatile conditions; and ii) model selection over the training period (7/2011–6/2017) being sub-optimal when applied to the volatile COVID-19 period. [14]

In summary, in real time the MS-VECM model successfully identified the regime switch at the beginning of the crisis. Noisy probabilities however suggest that a range of MS models should be used when trying to assess the presence of a switch in real time. MS-VECM forecasts provide better forecasts than the linear VECM model, however the differences are not statistically significant and may be a result of the small sample size.

5. Limitations and future research

While the MS-VECM models demonstrate improved prediction accuracy over linear benchmarks, these models have several practical limitations. First, the regime-switching framework introduces many additional parameters, particularly in the full MS-VECM. This increases the risk of overfitting and means that the model is generally unsuitable for small to modest sample sizes. Although the MSAI-VECM partially addresses this by restricting short-run dynamics, further research could explore dimension reduction techniques or Bayesian shrinkage to improve stability (Koop and Korobilis, 2009). Second, the analysis focuses on a relatively short out-of-sample forecast horizon of up to 12 months. Future work could evaluate model performance over longer horizons or under different forecast combinations and weighting schemes. Additionally, the forecasting exercise could be expanded to other non-linear models such factor-augmented VARs (Stock and Watson, 2002). Third, to identify regime switches, we required a long time span (25 years), which limited our ability to rigorously test out-of-sample performance. To reduce noise in regime probabilities, future research could also consider mixing student t (as opposed to Gaussian) distributions (Klaassen, 2002). Lastly, the MS-VECM assumes a state-invariant cointegrating vector. Long-run relationships between LTD, prices and volumes may vary across regimes, so regime-dependent cointegration may improve model fit and forecast precision (Rankin, 2005).

6. Conclusion

This paper has proposed the use of a Markov switching VECM (MS-VECM) when forecasting Land Transfer Duty (LTD) in Victoria. Our proposed model generally outperformed DTFs benchmark model, the linear VECM. The MS-VECM has plausible dynamics, identifying two states: a high-LTD-growth low-volatility state and a low-LTD-growth high-volatility state. The MS-VECM provides more accurate LTD forecasts over 3, 6, 9 and 12-month horizons. At the beginning of COVID-19, the model was also able to identify a regime switch (to the low-LTD-growth high-volatility state) in real time. When seeking to identify regime switches in real time though, we suggest estimating a range of MS specifications, as regime probabilities from this model class are notoriously noisy.

14 This latter issue could be overcome with a longer cross validation period, however we do not have enough observations to do so. Another possibility is the breakdown of long-term cointegrating relationships during the crisis. The assumption of a state-invariant cointegrating vector in our MS-VECM models may be false requiring state-dependent cointegration. This is left for further research.

References

Ang, A., & Bekaert, G. (2002). Regime switches in interest rates, *Journal of Business and Economic Statistics*, 20(2), 163–182.

Bai, J., & Perron, P. (2003). Computation and analysis of multiple structural change models, *Journal of Applied Econometrics*, 18(1), 1–22.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3), 307–327. https://doi.org/10.1016/0304-4076(86)90063-1

Chan, K.-S., & Tong, H. (1985). Testing linearity against smooth transition autoregressive models, *Biometrika*, 72(3), 491–499. https://doi.org/10.1093/ biomet/72.3.491

Cheung, Y.-W., & Lai, K. S. (1993). Finite-sample sizes of Johansen's likelihood ratio tests for cointegration, *Oxford Bulletin of Economics and Statistics*, 55(3), 313–328. https://doi.org/10.1111/j.1468-0084.1993.mp55003003.x

Clark, T. E. (2000). Forecasting an aggregate of cointegrated disaggregates, *Journal of Forecasting*, 19(1), 1–19.

Clements, M. P., & Hendry, D. F. (1999), *Forecasting non-stationary economic time series*. MIT Press.

CoreLogic Australia. (2023). *Property market update: July 2023,* [Accessed: 2023-08- 31]. https://www.corelogic.com.au

Dacco, R., & Satchell, S. (1999). Why do regime-switching models forecast so badly? *Journal of Forecasting*, 18(1), 1–16. https://doi. org/10.1002/(SICI)1099- 131X(199901)18:1⟨1::AID-FOR685⟩3.0.CO;2-B

Engle, R. F. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation, *Econometrica*, 50(4), 987–1007. https://doi.org/10.2307/1912773

Filardo, A. J. (1994). Business-cycle phases and their transitional dynamics. *Journal of Business and Economic Statistics*, 12(3), 299–308. https://doi.org/10. 2307/1392082

Guidolin, M. (2009). Regime shifts in empirical pricing kernels: A mixture capm. *CIREQ Montreal*.

Guidolin, M., Orlov, A. G., & Pedio, M. (2014). Understanding the impact of monetary policy shocks on the corporate bond market in good and bad times: A Markov switching model. *Proceedings of the European Financial Management Association Annual Meetings, Rome*.

Guidolin, M., & Timmermann, A. (2005). Economic implications of bull and bear regimes in UK stock and bond returns. *The Economic Journal*, 115(500), 111143.

Hamilton, J. D. (1989). A new approach to the economic analysis of non-stationary time series and the business cycle, *Econometrica*, 57(2), 357–384.

Hamilton, J. D. (1994). Time series analysis, *Princeton University Press*.

Hansen, P. R., Lunde, A., & Nason, J. M. (2011). The model confidence set, *Econometrica*, 79(2), 453–497.

Issam, D., Okba, A., & AbdelRahim, M. (2024). Structural changes in the impact of covid-19 pandemic on the performance of financial markets, Finan*cial Markets, Institutions and Risks*, 8, https://armgpublishing.com/journals/fmir/ volume-8-issue-2/ article-3/

Johansen, S. (1988). Statistical analysis of cointegration vectors, *Journal of Economic Dynamics and Control*, 12(2–3), 231–254.

Kim, C.-J. (1994). Dynamic linear models with Markov switching, *Journal of Econometrics*, 60, 1–22.

Kim, C.-J., & Nelson, C. R. (1999). *State-space models with regime switching: Classical and gibbs-sampling approaches with applications, MIT Press*.

Krolzig, H.-M. (1997a). A comparison of the forecast performance of Markov-switching and threshold autoregressive models of US GNP, *Oxford Bulletin of Economics and Statistics*, 59(4), 425–453. https://doi.org/10.1111/1468-0084.00080

Krolzig, H.-M. (1997b). *Markov-switching vector autoregressions: Modelling, statistical inference, and application to business cycle analysis* (Vol. 454), Springer-Verlag. https://doi.org/10.1007/978-3-642-51696-7

Krolzig, H.-M. (2001). Modelling the UK labour market with a Markov-switching vector equilibrium correction model, *The Economic Journal*, 111 (470), C215– C239. https://doi.org/10.1111/1468-0297.00643

Lim, G. C., & Tsiaplias, S. (2019). Forecasting with regime-switching models: New evidence from Australian GDP, *Economic Modelling*, 81, 52–70. https://doi. org/10.1016/j.econmod.2019.07.014

Psaradakis, Z., & Spagnolo, F. (2003). On the determination of the number of regimes in Markov-switching autoregressive models, *Journal of Time Series Analysis*, 24(2), 237–252. https://doi. org/10.1111/1467-9892.00310

Sarno, L., Valente, G., & Wohar, M. E. (2004). Monetary fundamentals and exchange rate dynamics under different nominal regimes. *Economic Inquiry*, 42(2), 179–191. https://doi.org/10.1093/ei/cbh053

Shen, X. (2014). *The stock market, the housing market, and consumption* [Doctoral dissertation, University of Waikato], https://researchcommons.waikato.ac. nz/handle/10289/8805

Sims, C. A. (1980). Macroeconomics and reality, *Econometrica*, 48 (1), 1–48.

State Revenue Office Victoria. (2024). *Land transfer duty (stamp duty)* [Accessed: 2024-09-14]. https://www.sro.vic.gov.au/land-transfer-duty

Trefz, N. M. (2023). S*tock price dynamics of us reits*, Springer. https://link.springer. com/content/pdf/10.1007/978-3-658-40049-1.pdf

Tsay, R. S. (1998). Testing and modeling threshold autoregressive processes, *Journal of the American Statistical Association*, 93(443), 1188–1202. https://doi.org/ 10.1080/01621459.1998.10473763

Appendix A

The AVG-n (10 per cent) models are simple forecast combinations that include models within the 10 per cent MCS $(M^*_{0.10})$ for the given n-month horizon, selected using training sample data from July 2011 to June 2017. Similarly, the AVG-n (25 per cent) models consist of models within the 25 per cent MCS $(M^*_{0.25})$ for the corresponding horizon, selected from the same training sample. The MS-Combo is a straight forward combination of the top-performing Markov switching models identified by the MCS. Notably, the confidence level and horizon do not influence the MS-Combo as the selected models remain consistent across cases. The models included in each combination are as follows:

- **AVG-3 (10 per cent)**: dVAR, VECM, VAR, MS-VAR, MS-VECM, MSAI-VECM
- **AVG-12 (10 per cent)**: VECM, MS-dVAR, MS-dAR, dAR, VAR, MSAI-VECM, MS-VAR, MS-VECM
- **AVG-3 (25 per cent)**: VECM, VAR, MS-VAR, MS-VECM, MSAI-VECM
- **AVG-12 (25 per cent)**: VAR, MSAI-VECM, MS-VAR, MS-VECM
- **MS-Combo**: MS-VECM, MSAI-VECM, MS-VAR

Section 4 presents the results for models chosen under CV(12), i.e. those that minimise the 12-month ahead MAPE over the training sample, as this is the primary forecast horizon of interest.

In contrast, the following tables display results for models selected under CV(3), where the selection criterion minimises the 3-month ahead MAPE over the training sample. Adjusting the forecast horizon for model selection can lead to different lag specifications, potentially affecting forecasting performance.

Table 8 summarises the lag specifications selected under CV(3) and compares them to the CV(12) models discussed in Section 4. Forecast evaluation results for the entire out-of-sample period and the COVID-19 crisis period, based on CV(3) models, are presented in Tables 9 and 10, respectively.

Table 8: Model summary

	LAGS (p)	
MODEL	CV(12)	CV(3)
VECM	p=4*	p=4*
VAR	p=2	p=12
dVAR	p=9	p=11
dAR	p=12	p=4
MS-VECM	p=3	p=2
MSAI-VECM	p=4	p=4
MS-VAR	p=2	p=4
MS-dVAR	p=4	p=1
MS-dAR	p=6	p=5

Note: $CV(n)$ represents the model lag specifications which minimise the n-step ahead cumulative MAPE over the training sample (7/2011-6/2017). Linear and MS models consider up to 12 and 6 lags, respectively. *Benchmark VECM lag specification is fixed at $p = 4$ (refer to footnote 13). Tables 6 and 7 report the CV(12) forecast results. Tables 9 and 10 report the CV(3) forecast results.

Table 9: CV(3) Models: Forecast evaluation (7/2017-6/2023)

	MAPE (%)				MSE (×10⁻¹⁶)			
HORIZON	**3**	**6**	**9**	**12**	**3**	**6**	**9**	**12**
VECM	7.34*	8.12	10.15*	11.53**	29.45*	134.12*	456.09**	1 092.85**
VAR	7.45**	8.35*	10.36*	12.16**	33.20*	144.48**	507.89**	1 263.89**
dVAR	7.53*	8.43	11.16	13.35	33.69*	154.02*	555.22*	1 357.17*
dAR	10.64	12.43	14.80	16.17*	52.68	291.77	975.11	2 277.82*
MS-VECM	6.97**	**7.30****	**8.68****	**10.45****	25.48**	**110.64****	**357.37****	**897.39****
MSAI-VECM	**6.77****	7.33**	9.74*	11.30**	**25.23****	113.26**	410.56**	1 021.85**
MS-VAR	7.31*	8.66	10.22	11.34**	31.28	151.34	479.99*	1 092.70**
MS-dVAR	11.41	14.13	16.94	20.12	72.48	431.81	1 474.82*	3 710.98*
MS-dAR	10.45	12.28	14.67	16.16*	48.93	275.14	946.06*	2 254.57*
AVG-3 (10%)	7.05**	7.94	9.89	11.41*	28.70	128.67	437.89*	1 056.34*
AVG-12 (10%)	7.22**	8.33**	10.54*	12.54**	29.07**	137.66**	479.85**	1 198.79**
AVG-3 (25%)	7.01**	7.92*	9.68*	11.09**	28.12	125.95*	423.19**	1 019.48**
AVG-12 (25%)	6.94**	7.87**	9.58*	11.00**	27.91	124.51*	416.79**	1 005.75**
MS-COMBO(3)	6.90**	7.71**	9.31*	10.67**	26.71**	120.39**	397.85**	955.36**

Note: The lowest loss for each horizon and loss function is underlined. **, * represent the model's inclusion in the MCS at 25 per cent ($M^*_{0.25}$) and 10 per cent ($M^*_{0.10}$)levels respectively.

Note that $M^*_{0.25} \subseteq M^*_{0.10}$. AVG-$n$ (10 per cent) and AVG-n (25 per cent) models refer to model averages for the n-horizon $M^*_{0.10}$ and $M^*_{0.25}$ determined during the training sample (7/2011-6/2017), respectively. MS-Combo is a simple average of the best performing Markov switching models. MSAI-VECM has state invariant autoregressive terms; all other parameters are state dependent.

Table 10: CV(3) Models: COVID-19 forecast evaluation (3/2020-8/2021)

	MSE (%)				MSE (×10⁻¹⁶)			
HORIZON	**3**	**6**	**9**	**12**	**3**	**6**	**9**	**12**
VECM	9.08**	12.51**	16.85	17.44	28.29**	204.62*	801.52*	1 449.29
VAR	8.95**	**11.43****	16.23*	17.20	28.75**	**172.07****	720.48**	1 402.11*
dVAR	10.00	12.91**	18.08**	20.18	36.57	230.42**	893.18**	1 785.46
dAR	14.15	17.41	16.35**	13.43**	70.52*	415.28	882.49*	1 092.13**
MS-VECM	8.54**	11.47**	15.72**	16.60	29.66*	210.01**	774.71**	1 245.81**
MSAI-VECM	8.06**	11.84**	16.83	17.76	**25.65****	191.87**	786.51	1 460.60
MS-VAR	8.40**	11.98**	15.59	16.84	30.21**	198.43**	710.62**	1 308.68
MS-dVAR	11.71**	12.49**	**13.60****	**12.89****	58.53**	294.13**	693.7**	**840.92****
MS-dAR	14.21	17.46	16.70**	14.02**	68.66	416.94	901.20*	1 141.71**
AVG-3 (10%)	8.47	11.95	16.56	17.43	28.57	197.27	766.83	1 404.26
AVG-12 (10%)	9.46**	12.24**	14.22**	14.30**	32.99**	216.95**	**673.72****	1 015.82**
AVG-3 (25%)	8.24**	11.78**	16.26	17.02	27.46**	192.19**	748.38**	1 345.89
AVG-12 (25%)	8.07**	11.59**	16.11	16.92	27.48**	189.90**	736.7**	1 322.87
MS-COMBO	**7.98****	11.64**	16.06*	17.01	27.81**	196.93**	749.35**	1 322.18

Note: The lowest loss for each horizon and loss function is underlined. **, * represent the model's inclusion in the MCS at 25 per cent $(M^*_{0.25})$ and 10 per cent $(M^*_{0.10})$levels respectively.

Note that $M^*_{0.25} \subseteq M^*_{0.10}$. AVG-$n$ (10 per cent) and AVG-n (25 per cent) models refer to model averages for the n-horizon $M^*_{0.10}$ and $M^*_{0.25}$ determined during the training sample (7/2011-6/2017), respectively. MS-Combo is a simple average of the best performing Markov switching models. MSAI-VECM has state invariant autoregressive terms; all other parameters are state dependent.