1 Discrete time framework for an Economic Scenario Generator (ESG)

1.1 VAR Model Construction

A Vector Autoregressive Model of order p, VAR(p) model can be used for jointly simulating various economic variables deriving the economy. The VAR(p) model can be represented as:

$$\mathbf{z}_t = \boldsymbol{\mu} + \Phi_1 \mathbf{z}_{t-1} + \Phi_2 \mathbf{z}_{t-2} + \cdots + \Phi_p \mathbf{z}_{t-p} + \boldsymbol{\epsilon},$$

where

- $\mathbf{z}_t$ is a vector of economic variables,
- $\boldsymbol{\mu}$ is a vector of intercepts,
- Φ_i, for $i = 1, \cdots, p$ are coefficient matrices of size $n \times n$ with n being the number of economic variables and p the lags.
- $\boldsymbol{\epsilon}$ is a vector of white noises.

Note that each variable is a linear combination of its past values and past values of other economic variables in the framework. We extend the framework presented in [1] to incorporate ten economic variables in table 1.

Variables	Definitions	Sources	Frequency
$r^{(1)}$	3-month zero-coupon yield	Reserve Bank of Australia	Daily
$r^{(40)}$	10-year zero-coupon yield	Reserve Bank of Australia	Daily
MR	Nominal mortgage rates	Reserve Bank of Australia	Monthly
HVI	Hedonic NSW home value index	CoreLogic	Monthly
y_t	Hedonic NSW rental yield	CoreLogic	Monthly
GDP	Australian nominal GDP	Australian Bureau of Statistics	Quarterly
CPI	Australia consumer price index	Australian Bureau of Statistics	Quarterly
ASX	S&P/ASX 200 Closing price	Refinitiv Eikon	Daily
UE	New South Wales Unemployment rate	Australian Bureau of Statistics	Monthly
AUD	Australian dollar trade-weighted index	Reserve Bank of Australia	Daily

Table 1: Notations, Definitions, Sources, and Frequency of Variables.

Data are available from April 1993 to March 2021. All data are transformed to quarterly frequency by averaging. In line with [1], we use the 3-month zero-coupon rate $r^{(1)}$ to proxy the level of the yield curve, and the 10-year zero-coupon spread $r^{(40)} - r^{(1)}$ to proxy the slope of the yield curve.

Different from [1], we use NSW houses instead of Sydney houses because the Sydney data from CoreLogic were variable, leading to extraordinarily large interquartile ranges (figure 1).

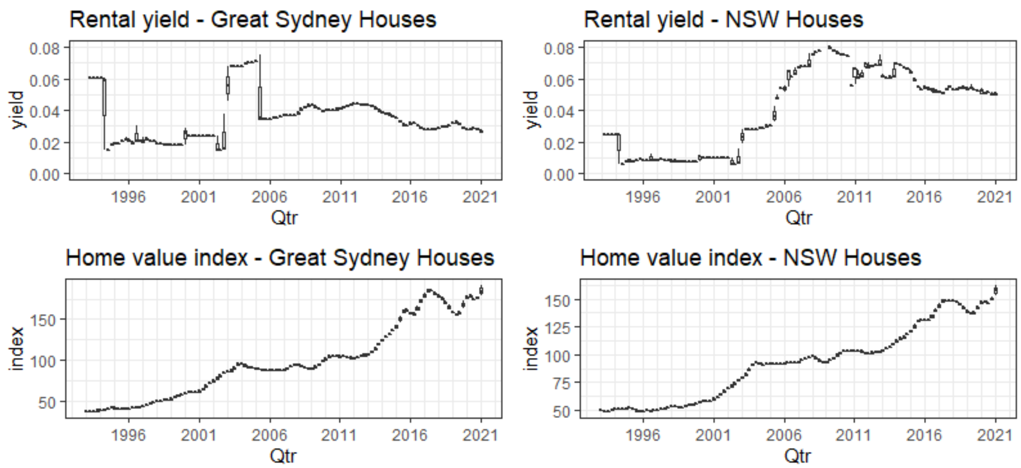

Figure 1: Comparisons of rental yields and home values for houses in Sydney and NSW.

Strong correlation can be found between the following two pairs, and the comparison can be found in figure 2.

- 3-month zero-coupon yield vs mortgage rate, which has a correlation coefficient of 90%. Mortgage rate will not be included in our VAR model and it can be simulated by adding 2.825% to the 3-month zero-coupon yield, which is the average margin between the two series from April 1993 to March 2021.
- 10-year zero-coupon spread vs NSW unemployment rate, which has a correlation coefficient of 75%. Similarly, NSW unemployment rate will be simulated by adding 4.956% to the 10-year zero-coupon spread trajectory. The relationship between these two variables has been examined in [4].

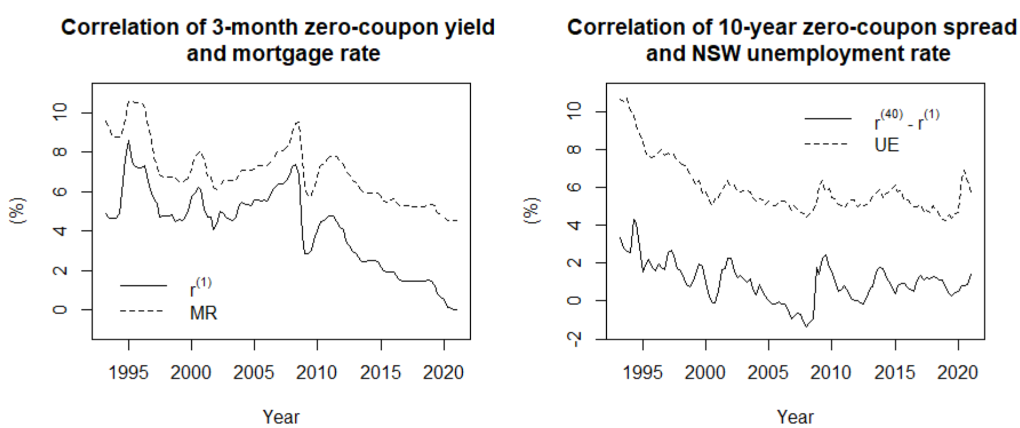

Figure 2: Comparisons of the two highly correlated pairs.

Further, we convert all indices to continuously compounding quarterly growth rates for consistency: $h_t = \log(HVI_t) - \log(HVI_{t-1})$, $g_t = \log(GDP_t) - \log(GDP_{t-1})$, $c_t = \log(CPI_t) - \log(CPI_{t-1})$, $a_t = \log(ASX_t) - \log(ASX_{t-1})$, and $d_t = \log(AUD_t) - \log(AUD_{t-1})$.

This leads to an eight-dimensional vector of economic variables:

$$\boldsymbol{z}_t = (r_t^{(1)}, r_t^{(40)} - r_t^{(1)}, h_t, y_t, g_t, c_t, a_t, d_t).$$

The plots for raw data and index growth rates are given in figures 3, 4.

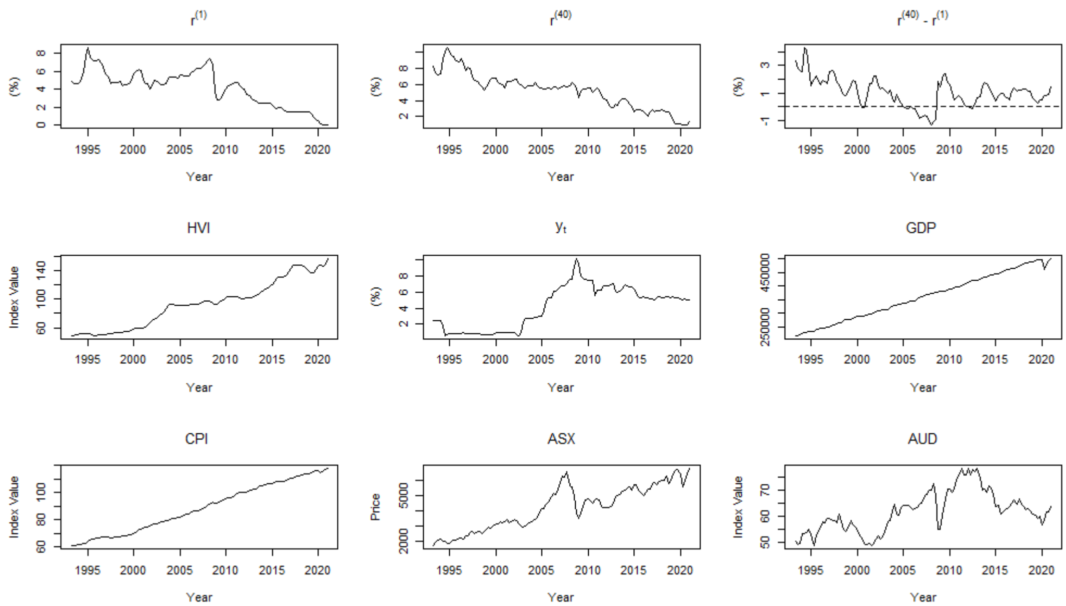

Figure 3: Raw data plots.

We can make interesting observations from the plots:

- Clear cyclical patterns can be spot out for the short-term interest rate $r^{(1)}$. As a response to the tight monetary policy after the 1990 recession [5], it peaked 8.6% in 1995. After a small hump around 2000 indicating the strong growth of the Australian economy [7], it climbed steadily to 7.4%, followed by a sharp plummet during the global financial crisis in 2008. As the economy recovered, it rose to 4.7% until 2011. It then fell and ultimately became negative during the COVID19 pandemic in 2021.

- The yield curve flattened gradually, and even became downward-sloping during the GFC. After that, the long term rate $r^{(40)}$ remained on average 1% higher than the short term rate $r^{(1)}$.

- NSW hedonic home value index doubled in the past decade. The rate of growth was particularly high when the interest rate was low [10].

- Investment in residential rental properties increased in early 1990 due to the 1987 stock market crash [2]. This was followed by a long quiet period. Rental yields are closed related to population growth which peaked in the late 2000s and slowed down thereafter[10] [3]. It leveled off at around 0.01 in the past 10 years.

- GDP and CPI increased at a constant rate before the COVID19 pandemic. One year after the pandemic shock, they recovered to the pre-pandemic levels of $501800 and 118 for GDP and CPI respectively.

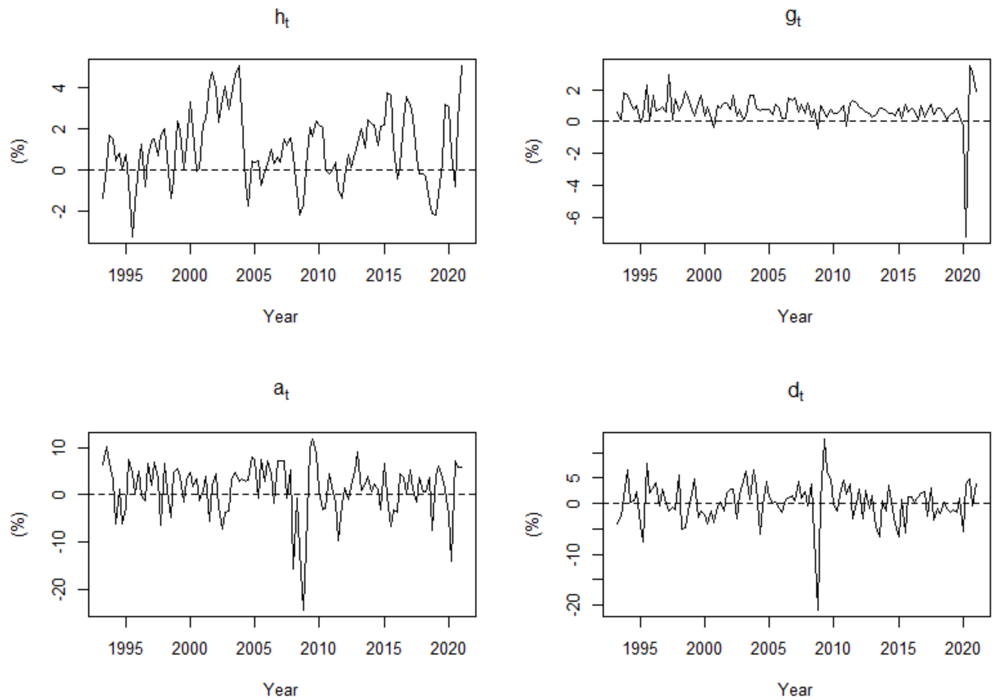

Figure 4: Transformed data plots.

- The S&P/ASX200 price grew steadily to \$6550 before the stock market crash in 2008. After small rise-and-falls, it touched the pre-GFC level again in 2021.
- The AUD trade-weighted index measures the price of Australian dollar against Australia's major trading partners [8]. The 1997 depreciation was due to the Asian recession [6]. During 2000-2013, despite a impacts of GFC, AUD increased as a result from the mining boom and inflation [11]. It then declined slowly which was driven by the decrease in the terms of trade and the ease of monetary policy [9].

1.2 Procedure for constructing a VAR model

We fit a VAR model with the eight variables described above.

1.2.1 Test for causality

The causal relationship among the variables are tested by Granger and instantaneous causality. The test statistics and p-values are reported in tables 2, 3. If X Granger-causes Y, past and current values of X forecast Y. If X instantaneously causes Y, past, current, and future values of X forecast Y. Under a significance level of 95%, we get the following Granger causal relationships:

- 3-month zero-coupon yield:

  - $r^{(1)} \to r^{(40)} - r^{(1)}$.
  - $r^{(1)} \to h_t$.
  - $r^{(1)} \to y_t$.
  - $r^{(1)} \to c_t$.
- 10-year zero-coupon spread:
  - $r^{(40)} - r^{(1)} \to r^{(1)}$.
  - $r^{(40)} - r^{(1)} \to y_t$.
- Home value index growth: does not Granger-cause anything.
- Rental yields: $y_t \to r^{(1)}$.
- GDP growth: $g_t \to h_t$.
- CPI growth:
  - $c_t \to a_t$.
  - $c_t \to d_t$.
- S&P/ASX200 price growth: $a_t \to r^{(1)}$.
- AUD trade-weighted index growth: does not Granger-cause anything.

And the following instantaneous causal relationships:

- 3-month zero-coupon yield:
  - $r^{(1)} \to r^{(40)} - r^{(1)}$.
  - $r^{(1)} \to y_t$.
  - $r^{(1)} \to g_t$.
  - $r^{(1)} \to c_t$.
  - $r^{(1)} \to a_t$.
  - $r^{(1)} \to d_t$.
- 10-year zero-coupon spread: $r^{(40)} - r^{(1)} \to r^{(1)}$.
- Home value index growth: $h_t \to y_t$.
- Rental yields:
  - $y_t \to r^{(1)}$.
  - $y_t \to h_t$.
  - $y_t \to a_t$.
  - $y_t \to d_t$.
- GDP growth:

- $g_t \to r^{(1)}$.
- $g_t \to c_t$.
- $g_t \to a_t$.

- CPI growth:
  - $c_t \to r^{(1)}$.
  - $c_t \to g_t$.
- S&P/ASX200 price growth:
  - $a_t \to r^{(1)}$.
  - $a_t \to y_t$.
  - $a_t \to g_t$.
  - $a_t \to d_t$.
- AUD trade-weighted index growth:
  - $d_t \to r^{(1)}$.
  - $d_t \to y_t$.
  - $d_t \to a_t$.

We can observe the following:

- There are more instantaneous causal relationships than Granger causal relationships.
- The 3-month zero-coupon yield has bidirectional relationships with most other variables, emphasising its importance in the VAR model.

	$r^{(1)}$		$r^{(40)} - r^{(1)}$		h_t		y_t		g_t		c_t		a_t		d_t	
	F	p	F	p	F	p	F	p	F	p	F	p	F	p	F	p
$r^{(1)}$	.	.	4.203	0.016	4.249	0.006	3.252	0.041	2.309	0.078	4.844	0.003	0.522	0.667	2.041	0.132
$r^{(40)} - r^{(1)}$	10.270	0.000	.	.	2.150	0.095	8.170	0.000	1.359	0.259	0.936	0.334	1.163	0.282	1.856	0.159
h_t	1.228	0.301	1.053	0.370	.	.	0.088	0.967	1.014	0.387	0.125	0.945	0.507	0.678	1.232	0.299
y_t	6.599	0.002	0.316	0.729	0.083	0.969	.	.	1.702	0.185	1.450	0.237	1.178	0.310	1.157	0.316
g_t	0.786	0.503	0.301	0.740	3.993	0.009	0.149	0.862	.	.	2.057	0.153	0.142	0.707	0.001	0.981
c_t	1.172	0.322	0.049	0.825	0.539	0.656	2.012	0.136	1.001	0.318	.	.	5.485	0.020	4.455	0.036
a_t	3.164	0.026	0.582	0.446	0.945	0.420	0.964	0.383	0.982	0.323	0.470	0.494	.	.	0.043	0.836
d_t	1.792	0.169	0.315	0.730	0.197	0.898	1.649	0.195	1.666	0.198	1.516	0.219	0.377	0.540	.	.

Table 2: F-statistics and p-values for the Granger causality test.

	$r^{(1)}$		$r^{(40)} - r^{(1)}$		h_t		y_t		g_t		c_t		a_t		d_t	
	χ^2	p	χ^2	p	χ^2	p	χ^2	p	χ^2	p	χ^2	p	χ^2	p	χ^2	p
$r^{(1)}$	.	.	24.057	0.000	0.587	0.443	4.528	0.033	7.658	0.006	6.256	0.012	4.831	0.028	25.301	0.000
$r^{(40)} - r^{(1)}$	24.057	0.000	.	.	0.002	0.964	2.018	0.155	0.478	0.490	2.212	0.137	1.061	0.303	3.685	0.055
h_t	0.587	0.443	0.002	0.964	.	.	4.172	0.041	0.058	0.810	0.964	0.326	2.213	0.137	2.613	0.106
y_t	4.528	0.033	2.018	0.155	4.172	0.041	.	.	0.383	0.536	3.008	0.083	5.737	0.017	5.618	0.018
g_t	7.658	0.006	0.478	0.490	0.058	0.810	0.383	0.536	.	.	9.082	0.003	7.526	0.006	0.382	0.537
c_t	6.256	0.012	2.212	0.137	0.964	0.326	3.008	0.083	9.082	0.003	.	.	3.055	0.081	0.425	0.515
a_t	4.830	0.028	1.061	0.303	2.213	0.137	5.737	0.017	7.526	0.006	3.055	0.081	.	.	17.666	0.000
d_t	25.301	0.000	3.685	0.055	2.613	0.106	5.618	0.018	0.382	0.537	0.425	0.515	17.666	0.000	.	.

Table 3: Chi-squared statistics and p-values for the Instantaneous causality test.

1.2.2 Test for stationarity

We conduct the Augmented Dickey–Fuller (ADF) test and Phillips-Perron (PP) test for stationarity of all eight series. The T-statistics and the corresponding p-values are tabulated in table 4. Under a significant level of 5%, we are not confident that $r^{(1)}$ and y_t are stationary.

	ADF		PP	
Variables	t-statistic	p-value	t-statistic	p-value
$r^{(1)}$	-1.0581	0.2990	-1.46344	0.4485
$r^{(40)} - r^{(1)}$	-2.4485	0.0163	-9.23135	0.0357
h_t	-3.28511	<0.01	-21.2809	<0.01
y_t	-0.09205	0.6166	-0.18647	0.6476
g_t	-7.34669	<0.01	-83.6129	<0.01
c_t	-5.41503	<0.01	-47.4448	<0.01
a_t	-7.44914	<0.01	-68.8355	<0.01
d_t	-8.73256	<0.01	-77.771	<0.01

Table 4: Stationary test statistics before transformations.

From figure 3, we notice clear cyclical patterns with a decreasing trend for $r^{(1)}$, and slow increasing trend for y_t. Thus, we take the first difference of them, which make them stationary (table 5). Now, all series now start from July 1993, totally 112 quarters. Denote them

$$\begin{aligned} z_t &= \left(r_t^{(1)} - r_{t-1}^{(1)}, r_t^{(40)} - r_t^{(1)}, h_t, y_t - y_{t-1}, g_t, c_t, a_t, d_t\right)^\top \\ &= \left(\tilde{r}_t^{(1)}, r_t^{(40)} - r_t^{(1)}, h_t, \tilde{y}_t, g_t, c_t, a_t, d_t\right)^\top. \end{aligned}$$

	ADF		PP	
Variables	t-statistic	p-value	t-statistic	p-value
$\tilde{r}^{(1)}$	-6.42009	<0.01	-55.1799	<0.01
$\tilde{y}_t$	-8.16077	<0.01	-79.6425	<0.01

Table 5: Stationary test statistics after differencing.

1.2.3 Find the optimal lag order

We compare the Akaike Information Criterion (AIC), Schwarz Information Criterion (SIC), and Hannan-Quinn Criterion (HQC) up to lag 10, to find the optimal lag p. From table 6, all information criteria agree on lag 1, which is a very parsimonious model.

Lag Order	AIC	SIC	HQC
1	-73.95640 *	-72.38520 *	-73.31910 *
2	-72.83040	-69.66990	-71.54870
3	-71.61260	-66.84440	-69.67930
4	-70.67850	-64.28370	-68.08610
5	-69.77340	-61.73280	-66.51450
6	-68.71220	-59.00630	-64.77920
7	-67.77340	-56.38220	-63.15850
8	-67.46120	-54.36430	-62.15650
9	-67.02290	-52.19960	-61.02050
10	-67.86050	-51.28940	-61.15210

Table 6: Lag order selection, with the optimal orders marked by asterisks.

1.2.4 Train the model

The most recent 20% observations are left out as the testing set. We build a VAR(1) model on the training set.

1.2.5 Evaluate the model

We compare the forecasts against the actual data on figure 5.

- The severe disruptions from COVID-19 is not modelled by VAR(1).
- The overall trends for the 3-month zero-coupon yield/spread, home value index, GDP, S&P/ASX200 are correctly identified.
- The 95% confidence bounds are narrower for series that are less volatile, such as GDP and CPI.
- VAR(1) significantly overestimated rental yields and AUD trade-weighted index, which could lead to overoptimistic investment returns forecast.

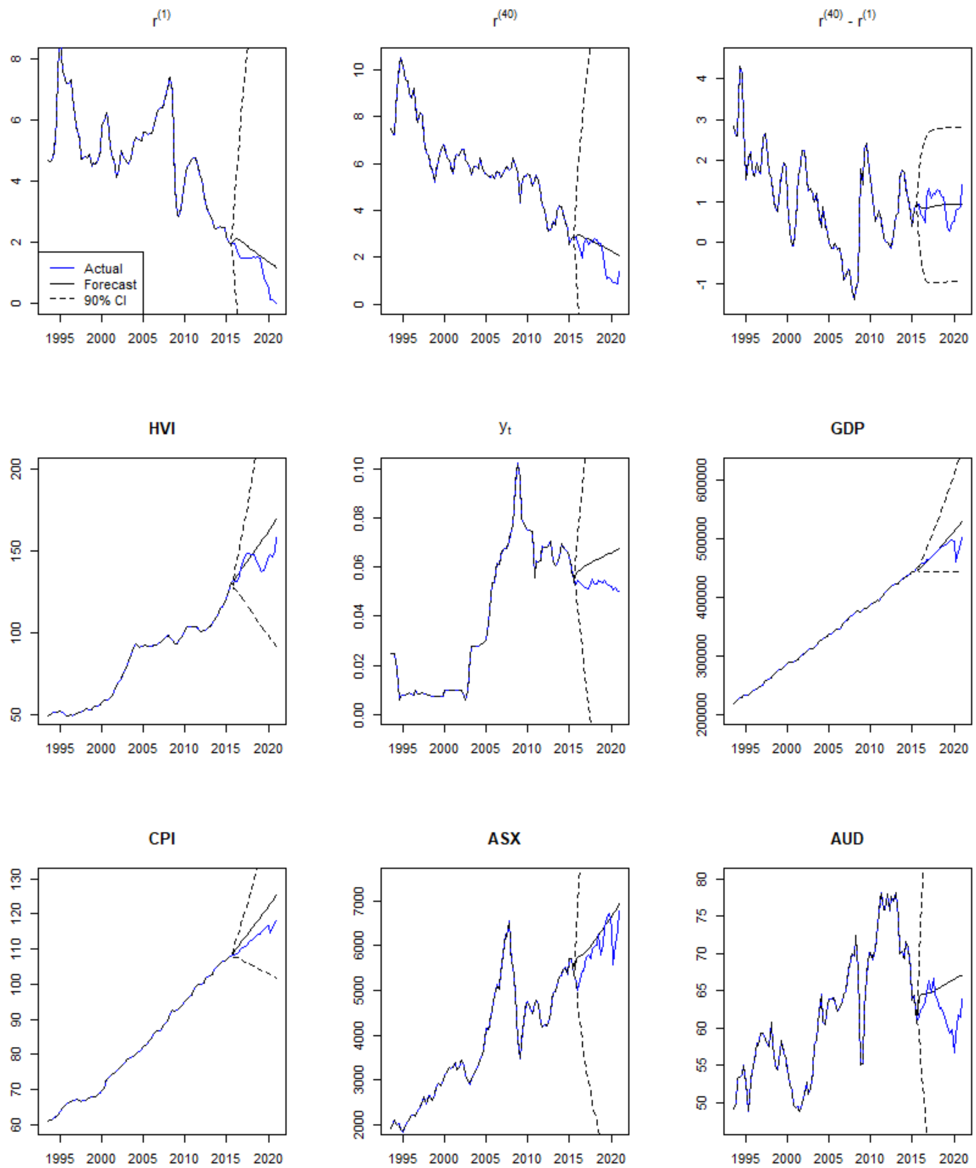

Figure 5: Forecasts vs actual on the original scale.

Table 7 reports the root mean squared errors (RMSE) and mean absolute errors (MAE).

	$\tilde{r}^{(1)}$	$r^{(40)} - r^{(1)}$	h_t	$\tilde{y}_t$	g_t	c_t	a_t	d_t	overall
RMSE	0.00139	0.00340	0.02079	0.00123	0.01926	0.00680	0.05014	0.02506	0.02238
MAE	0.00099	0.00302	0.01878	0.00112	0.00899	0.00401	0.03818	0.02094	0.01200

Table 7: Overall RMSE and MAE and breakdowns for the VAR(1) model.

	μ		Φ															
	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE	Coeff	SE
$\tilde{r}^{(1)}$	-0.002	0.001	0.340	0.092	0.114	0.039	0.029	0.023	-0.196	0.095	-0.021	0.040	0.004	0.072	0.010	0.008	0.007	0.011
$r^{(40)}$	0.000	0.001	-0.451	0.110	0.867	0.047	0.022	0.027	0.020	0.114	0.044	0.048	0.057	0.087	-0.006	0.009	0.019	0.014
h_t	0.001	0.002	-0.662	0.286	0.057	0.121	0.726	0.071	0.075	0.297	0.309	0.125	-0.177	0.225	0.031	0.023	0.010	0.035
$\tilde{y}_t$	0.001	0.001	0.150	0.098	-0.107	0.041	-0.002	0.024	0.197	0.102	-0.034	0.043	0.048	0.077	0.008	0.008	0.001	0.012
g_t	0.007	0.002	-0.243	0.238	0.174	0.101	-0.066	0.059	0.309	0.247	-0.062	0.104	-0.151	0.188	0.018	0.019	0.044	0.029
c_t	0.007	0.001	0.257	0.133	-0.075	0.056	0.006	0.033	-0.186	0.138	-0.091	0.058	0.088	0.105	0.004	0.011	-0.002	0.016
a_t	0.021	0.011	-0.168	1.299	0.125	0.551	-0.325	0.323	-1.153	1.349	0.132	0.567	-1.931	1.025	0.270	0.106	0.126	0.160
d_t	-0.002	0.007	-2.216	0.910	0.349	0.386	0.264	0.226	-1.121	0.945	0.237	0.397	-0.778	0.718	-0.011	0.074	0.240	0.112

Table 8: Estimated coefficients and standard errors of the VAR(1) model on the whole data set.

1.2.6 Diagnostics on the full dataset

We fit a VAR(1) on the full dataset:

$$\tilde{\mathbf{z}}_t = \boldsymbol{\mu} + \tilde{\Phi}\tilde{\mathbf{z}}_{t-1} + \Sigma^{1/2}\boldsymbol{\epsilon}$$

where $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, I)$.

The coefficients and standard errors are presented in table 8, and the covariance/correlation matrices are in tables 9, 10.

	$\tilde{r}^{(1)}$	$r^{(1)} - r^{(40)}$	h_t	$\tilde{y}_t$	g_t	c_t	a_t	d_t
$\tilde{r}^{(1)}$	1.48E-05	-1.01E-05	2.53E-06	-2.71E-06	6.60E-06	4.41E-06	4.33E-05	7.57E-05
$r^{(40)}$	-1.01E-05	2.12E-05	-6.24E-07	-1.21E-06	-4.37E-06	-1.45E-06	-2.66E-05	-4.79E-05
h_t	2.53E-06	-6.24E-07	1.43E-04	-6.83E-06	8.61E-06	-1.01E-06	1.24E-04	4.30E-05
y_t	-2.71E-06	-1.21E-06	-6.83E-06	1.68E-05	-3.35E-08	2.47E-06	-4.81E-05	-2.81E-05
g_t	6.60E-06	-4.37E-06	8.61E-06	-3.35E-08	9.90E-05	2.00E-05	1.44E-04	7.73E-06
c_t	4.41E-06	-1.45E-06	-1.01E-06	2.47E-06	2.00E-05	3.09E-05	5.36E-05	2.51E-05
a_t	4.33E-05	-2.66E-05	1.24E-04	-4.81E-05	1.44E-04	5.36E-05	2.96E-03	9.07E-04
d_t	7.57E-05	-4.79E-05	4.30E-05	-2.81E-05	7.73E-06	2.51E-05	9.07E-04	1.45E-03

Table 9: Covariance matrix Σ of the residuals on the full dataset.

	$\tilde{r}^{(1)}$	$r^{(1)} - r^{(40)}$	h_t	$\tilde{y}_t$	g_t	c_t	a_t	d_t
$\tilde{r}^{(1)}$	1.000	-0.571	0.055	-0.172	0.173	0.206	0.207	0.517
$r^{(40)} - r^{(1)}$	-0.571	1.000	-0.011	-0.064	-0.095	-0.056	-0.106	-0.273
h_t	0.055	-0.011	1.000	-0.139	0.072	-0.015	0.190	0.094
$\tilde{y}_t$	-0.172	-0.064	-0.139	1.000	-0.001	0.108	-0.216	-0.180
g_t	0.173	-0.095	0.072	-0.001	1.000	0.362	0.265	0.020
c_t	0.206	-0.056	-0.015	0.108	0.362	1.000	0.177	0.118
a_t	0.207	-0.106	0.190	-0.216	0.265	0.177	1.000	0.438
d_t	0.517	-0.273	0.094	-0.180	0.020	0.118	0.438	1.000

Table 10: Correlation matrix of the residuals on the full dataset.

We conduct the Portmanteau test, Engle's ARCH test for residual autocorrelation, as well as the Jarque-Bera test for residual normality. The Chi-squared statistics and the corresponding pvalues are summarised in table 11. Unfortunately, the residuals of this VAR(1) model are not white noise, contradicting the model assumptions.

	Null hypotheses	χ^2 statistics	p-values
Portmanteau test	No autoccorelation	31.1431	0.0000
Engle's ARCH test	No autoccorelation	1512.8429	0.0000
Jarque-Bera test	Same skewness as the Gaussian distribution	478.9223	0.0000
	Same kurtosis as the Gaussian distribution	5996.4861	0.0000

Table 11: Residual test statistics and p-values.

From the residual plots (figure 6), we notice:

- Most residuals look random, have mean 0 and constant variance.
- The model fails to capture the Global Financial Crisis trauma. This is obvious for the interest rate yield curve, rental yields, S&P/ASX200 index, and AUD trade weighted-index.
- COVID19 impacts are more moderate.
- CPI has large residuals around 2000, when Australia was experiencing strong growth [7], as well as at 2020 during COVID19.

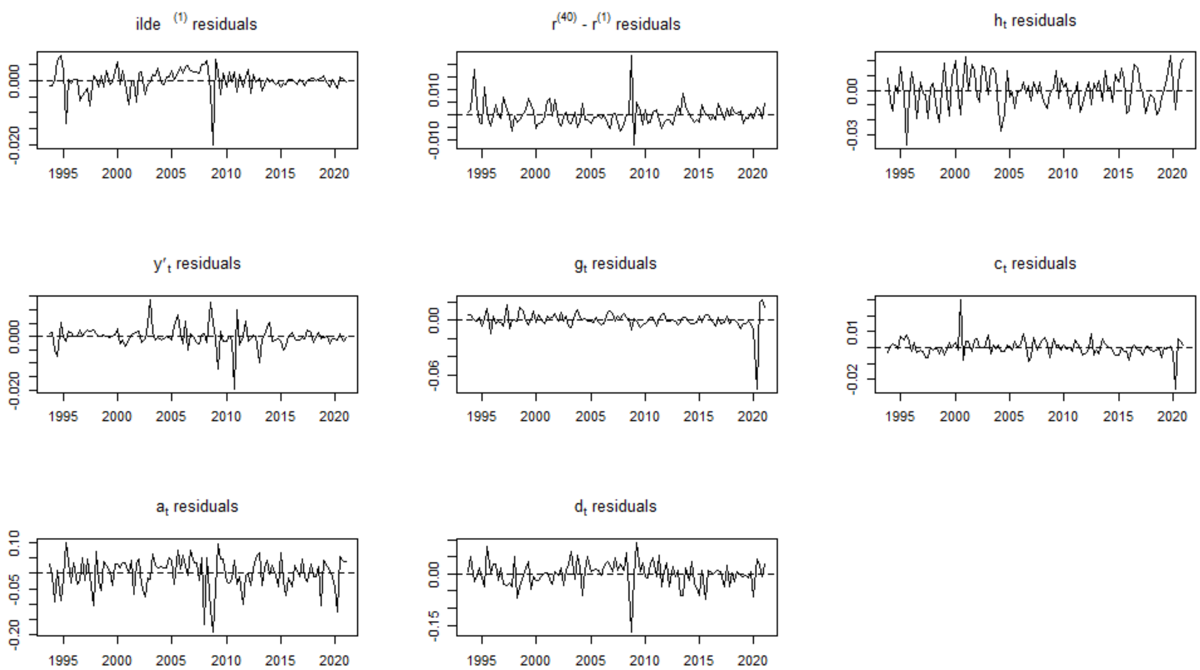

Figure 6: Residual plots.

Reading from the Normal QQ-plots (figure 7):

- All series are normal in the middle.
- Rental yield series has heavier tails than the normal distribution, while the 3-month zero-coupon rate and S&P/ASX200 are left-skewed, 10-year zero-coupon spread is right-skewed.
- CPI and AUD trade-weighted index have outliers.

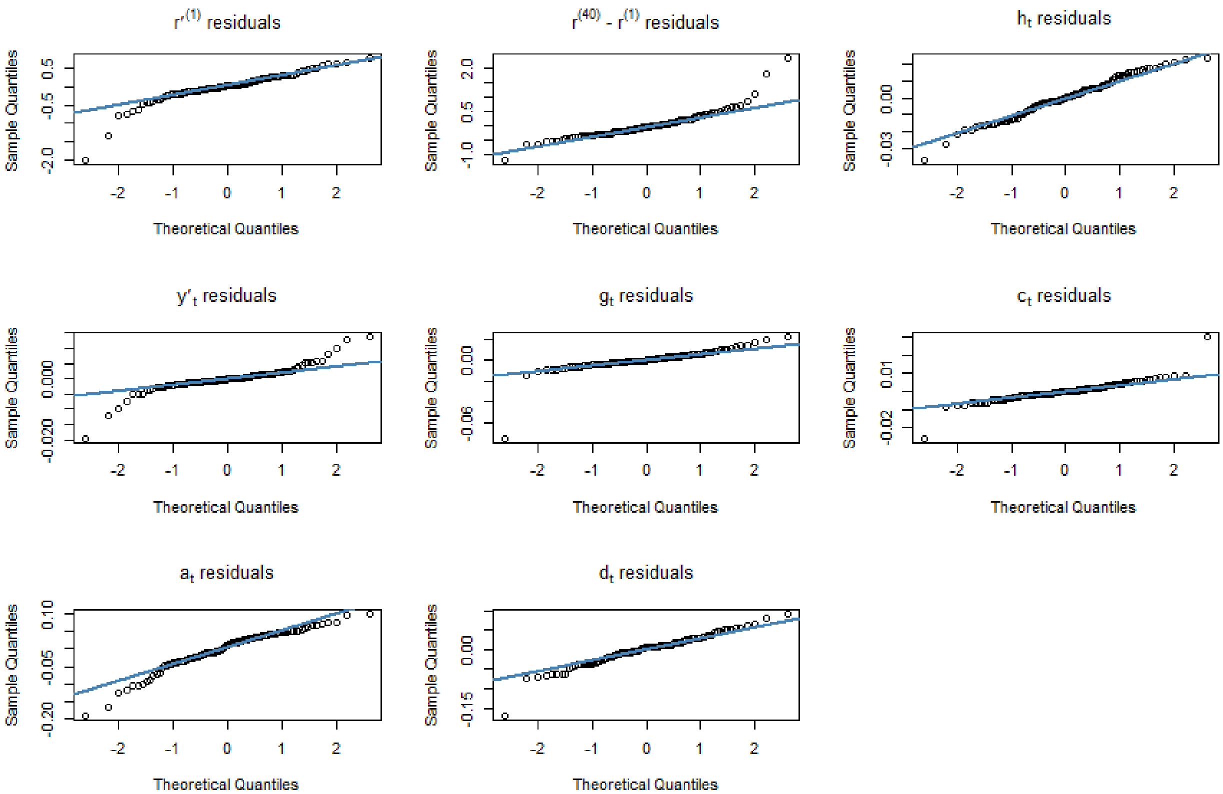

Figure 7: Normal QQ-plots of the residuals.

The QQ plots above do not reveal the multivariate (non-)normality.

1.2.7 Forecast

We simulated 10,000 trajectories for 10 years, and forecasts on the original scale and on the transformed scale are shown in figure 8, 9. Some problems should be noted:

- The prediction stablises after 7-8 years. That is, the series mean continue to grow/decay at the same direction and rate. A reason is that a VAR model with 25 years of historic data cannot support long forecasting periods. **(I would suggest at most 10 years.)** This phenomenon introduces some undesirable features because most economic variables have natural bounds. For example, interest rate can be negative, but it is not likely to be -5% or -10%.
- The 90% confidence intervals span most reasonable values on the transformed scale (figure 9).
- Judgements are required in case of extreme events.

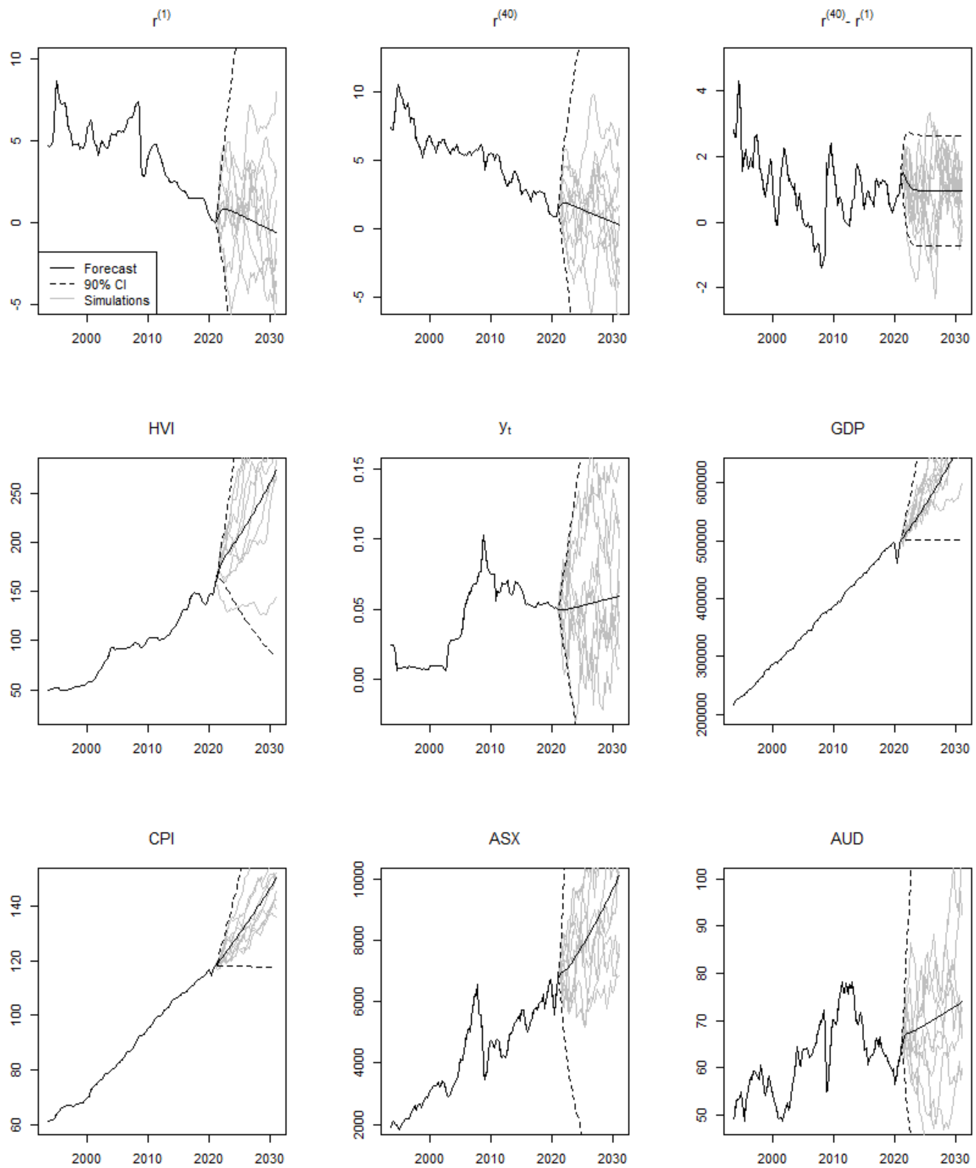

Figure 8: 10-year forecasts on the original scale.

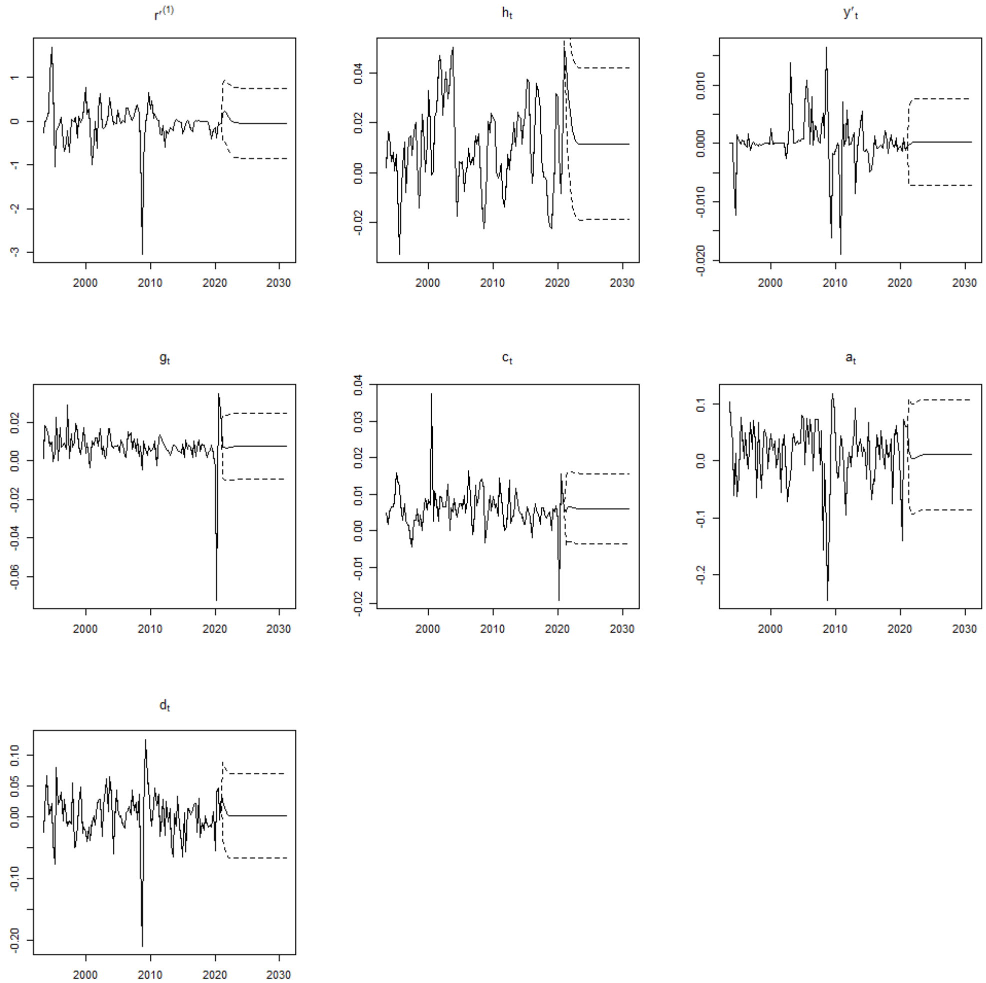

Figure 9: 10-year forecasts on the transformed scale.

Figure 10 demonstrates how the historic empirical cumulative distribution functions (ECDFs) are comparable to the simulated ones.

- The historic and simulated distributions for 10-year zero-coupon spread, S&P/ASX200 growth, and AUD trade-weighted index growth are almost identical.
- The simulated distribution for the remaining variables are more spread out, though their medians are similar.
- We note that the less spikes there are in the residuals (relatively), the closer the ECDFs.

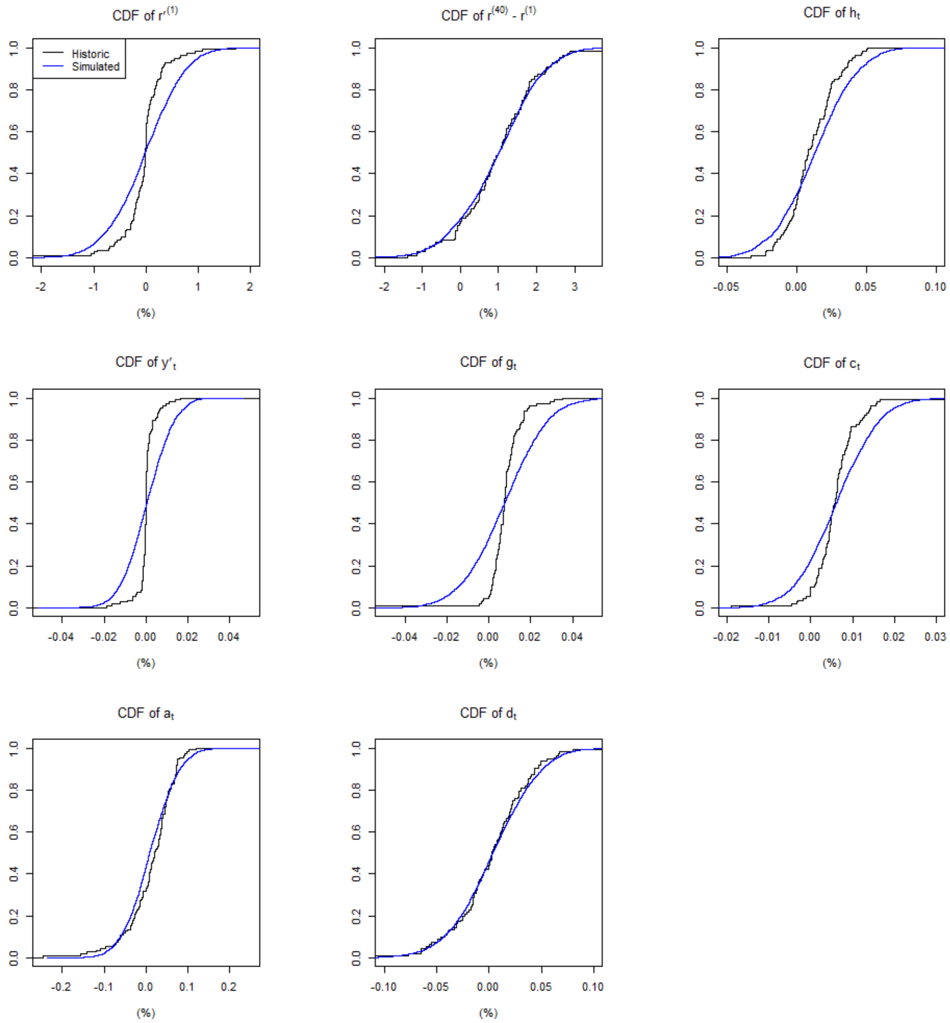

Figure 10: ECDFs of historic data vs 10-year forecasts on the transformed scale.

1.3 A Cascade Model

1.3.1 Structure

The first level is a VAR model containing the stationary series $\tilde{r}^{(1)}, r^{(40)} - r^{(1)}, h_t, \tilde{y}_t, g_t$. The second level then regress the upper level variables on each of the remaining variables. This structure, visualised in figure 11, is consistent with [1] as well as the causal relationships identified above.

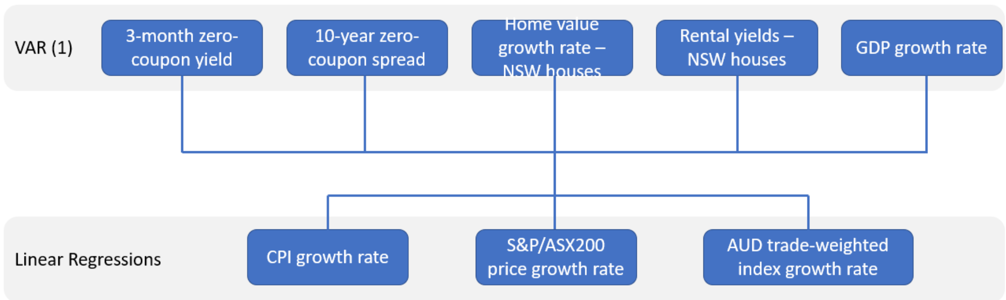

Figure 11: The cascade structure.

1.3.2 Procedure for constructing a VAR model for the first level

The optimal lag order for the first level is still 1 (table 12).

Lag Order	AIC	SIC	HQC
1	-50.9815 *	-50.3678 *	-50.7326 *
2	-50.5928	-49.3583	-50.0922
3	-50.1467	-48.2841	-49.3915
4	-49.8936	-47.3957	-48.881
5	-49.4974	-46.3566	-48.2244
6	-48.9841	-45.1927	-47.4477

Table 12: Lag order selection, with the optimal orders marked by asterisks.

On the same training set, we build a VAR(1):

$$\tilde{\mathbf{z}}_t = \tilde{\boldsymbol{\mu}} + \tilde{\Phi}_1 \tilde{\mathbf{z}}_{t-1} + \tilde{\Sigma}^{1/2} \boldsymbol{\epsilon},$$

where $\tilde{\mathbf{z}}_t = (\tilde{r}^{(1)}, r^{(40)} - r^{(1)}, h_t, \tilde{y}_t, g_t)$. The coefficients and standard errors are presented in table 13, and the covariance and correlation matrices are in table 14.

	μ		Φ									
$\tilde{r}^{(40)}$	-0.002	0.001	0.384	0.082	0.110	0.039	0.032	0.023	-0.229	0.092	-0.006	0.037
$r^{(40)} - r^{(1)}$	0.000	0.001	-0.384	0.097	0.863	0.046	0.024	0.027	0.015	0.109	0.046	0.045
h_t	0.000	0.002	-0.620	0.254	0.061	0.120	0.737	0.070	-0.053	0.285	0.328	0.116
$\tilde{y}_t$	0.002	0.001	0.187	0.086	-0.112	0.041	-0.001	0.024	0.186	0.097	-0.015	0.040
g_t	0.007	0.002	-0.108	0.214	0.177	0.101	-0.049	0.059	0.162	0.240	-0.058	0.098

Table 13: Estimated coefficients and standard errors of the first level VAR(1) model on the training set.

1.3.3 Procedure for constructing linear regressions for the second level

We regress the variables in level 1 on each of the remaining variables c_t, a_t, d_t. Compared to a VAR model which only incorporates past values, linear regressions include both past

$\tilde{\Sigma}$ (5 × 5)					$\tilde{\rho}$ (5 × 5)				
1.48E-05	-9.70E-06	3.58E-06	-2.33E-06	7.60E-06	1.000	-0.550	0.078	-0.149	0.197
-9.70E-06	2.10E-05	-7.40E-07	-1.10E-06	-3.64E-06	-0.550	1.000	-0.013	-0.059	-0.079
3.58E-06	-7.40E-07	1.43E-04	-6.04E-06	1.21E-05	0.078	-0.013	1.000	-0.124	0.101
-2.33E-06	-1.10E-06	-6.04E-06	1.66E-05	5.02E-07	-0.149	-0.059	-0.124	1.000	0.012
7.60E-06	-3.64E-06	1.21E-05	5.02E-07	1.01E-04	0.197	-0.079	0.101	0.012	1.000

Table 14: Covariance matrix $\tilde{\Sigma}$ (left) and correlation matrix $\tilde{\rho}$ (right) of the residuals on the training set.

and current values of the regressors:

$$
\begin{aligned}
c_t &= \alpha + \beta_1 r_t^{\prime(1)} + \beta_2 (r_t^{(40)} - r_t^{(1)}) + \beta_3 h_t + \beta_4 y_t^{\prime} + \beta_5 g_t + \epsilon, \\
a_t &= \alpha + \beta_1 r_t^{\prime(1)} + \beta_2 (r_t^{(40)} - r_t^{(1)}) + \beta_3 h_t + \beta_4 y_t^{\prime} + \beta_5 g_t + \epsilon, \\
d_t &= \alpha + \beta_1 r_t^{\prime(1)} + \beta_2 (r_t^{(40)} - r_t^{(1)}) + \beta_3 h_t + \beta_4 y_t^{\prime} + \beta_5 g_t + \epsilon.
\end{aligned}
$$

The estimated coefficient are presented in table 15. No variable is significant according to the T tests. **Note: I didn't fit this cascade structure on the full dataset because it has larger errors than the VAR(1) model with all variables. The estimated coefficients/standard errors are on the training set only.**

	c_t		a_t		d_t	
	Coeff	SE	Coeff	SE	Coeff	SE
(Intercept)	0.0095	0.0011	0.0128	0.0126	-0.0050	0.0083
$\tilde{r}^{(1)}$	0.3819	0.1083	2.3919	1.2330	3.1646	0.0081
$r^{(40)} - r^{(1)}$	-0.0517	0.0513	-0.1547	0.5843	-0.1268	0.0038
h_t	-0.0372	0.0339	0.2666	0.3866	0.3043	0.2545
$\tilde{y}_t$	0.1614	0.1210	-2.4226	1.3784	-1.6317	0.9074
g_t	-0.2477	0.0957	0.0285	1.0901	0.8247	0.7176

Table 15: Estimated coefficients and standard errors of the second level linear regressions on the training set.

1.3.4 Evaluating the cascade model on the training set

The same residual tests are performed for the first level. Table 16 indicates that the residuals are not white noise.

	Null hypotheses	χ^2 statistics	p-values
Portmanteau test	No autocorrelation	11.9630	0.0000
Engle's ARCH test	No autocorrelation	316.8227	0.0001
Jarque-Bera test	Same skewness as the Gaussian distribution	92.1256	0.0000
	Same kurtosis as the Gaussian distribution	350.5271	0.0000

Table 16: Residual test statistics and p-values of the first level VAR(1) model.

Table 17 shows that the RMSE and MAE for this cascade structure are slightly higher than the VAR model with all eight variables.

	$\tilde{r}^{(1)}$	$r^{(1)} - r^{(40)}$	h_t	$\tilde{y}_t$	g_t	c_t	a_t	d_t	overall
RMSE	0.001527	0.003466	0.02097	0.001293	0.019283	0.006495	0.051194	0.02448	0.02261
MAE	0.001114	0.003098	0.018911	0.001198	0.009057	0.003642	0.039036	0.019486	0.01194

Table 17: Overall RMSE and MAE and breakdowns for the cascade model.

1.3.5 Forecasts

The cascade model doesn't outperform the VAR(1) model with all eight variables in terms of RMSE, and it only performs slightly in terms of MAE. So I did not do forecasts.

References

[1] Daniel H Alai, Hua Chen, Daniel Cho, Katja Hanewald, and Michael Sherris. Developing equity release markets: Risk analysis for reverse mortgages and home reversions. *North American Actuarial Journal*, 18(1):217–241, 2014.

[2] Australian Bureau of Statistics. Housing Stock: Investment in residential rental property, Jun 1995.

[3] Australian Bureau of Statistics. FEATURE ARTICLE: THE AUSTRALIAN RESIDENTIAL PROPERTY MARKET, Dec 2015.

[4] Zvi Eckstein, Ofer Setty, and David Weiss. Financial risk and unemployment. *International Economic Review*, 60(2):475–516, 2019.

[5] Reserve Bank of Australia. Statement by the Governor, Mr Bernie Fraser: Monetary Policy to Tighten: Media Releases, Aug 1994.

[6] Reserve Bank of Australia. Reserve Bank Bulletin – August 1998: The Economy and Financial Markets, Aug 1998.

[7] Reserve Bank of Australia. Article 1: 1999-2000 in review: Australia's strong growth exceeds expectations, Dec 2000.

[8] Reserve Bank of Australia. Weights for the Australian dollar trade-weighted index, Nov 2020.

[9] Reserve Bank of Australia. Bulletin – March 2021: Determinants of the Australian Dollar Over Recent Years, Mar 2021.

[10] Trent Saunders and Peter Tulip. A model of the australian housing market. *Economic Record*, 96:1–25, 2020.

[11] Treasury of the Australian Government. Understanding the appreciation of the Australian dollar and its policy implications, Mar 2012.