# Yet More on a Stochastic Economic Model: Part 5: A VAR model for Retail Prices and Wages

A. D. Wilkie
*InQA Limited, Dennington, Ridgeway, Horsell, Woking, GU21 4QR, U. K.;*
*email: david.wilkie@inqa.com.*

Şule Şahin
*Institute for Financial and Actuarial Mathematics, University of Liverpool, Mathematical Sciences Building, Liverpool L69 7ZL, UK*
*and Department of Actuarial Sciences, Hacettepe University, 06800, Ankara, Turkey;*
*e-mail: Sule.Sahin@liverpool.ac.uk*

**Revised January 2018**

**Abstract**

In this paper we develop a vector autoregressive (VAR) model for retail prices and wages within the Wilkie model. The results turn out to be a slight improvement over the original model, but the simulated results are not very different.

**Keywords**

Wilkie model; retail prices; wages; vector autoregressive (VAR) model

## 1. Introduction

1.1 This paper is the next in a series that updates the Wilkie model, originally put forward in Wilkie (1986) and Wilkie (1995), and follows Part 1 (Wilkie et al, 2011), Part 2 (Wilkie & Şahin, 2016), Parts 3A, 3B and 3C (Wilkie & Şahin, 2017a, 2017b and 2017c) and Part 4 (Wilkie & Şahin, 2018). In Part 4 we observed that in the model for retail prices there was a correlation with wages, which we had not allowed for, between the residuals of the inflation rate model in any year, and the residuals of the wages model in the previous year, which suggested that a vector autoregressive (VAR) model might be appropriate. We investigate this suggestion in this paper.

1.2 The first reason for having modelled the retail prices index, *Q*, without reference to the wages index, *W*, was that in the first version of the model in Wilkie (1986), a wages index was not included, being of less immediate importance than those variables that were modelled. It was included in Wilkie (1996), because it was thought to be necessary for the valuation of pension schemes with salary-related benefits, and possibly for assessing future liability claims in general insurance and in future wage and salary expenses in any organisation. These advantages still apply.

1.3 However, some users of stochastic models (now often called Economic Scenario Generators) may consider that wages are an unnecessary complication of the model, and making them essential, as we do in this paper, is an undesirable feature.

1.4 Another consideration is that good international figures for average earnings are not readily available. Some countries seem to prefer to publish indices of hourly wage rates in manufacturing industry, which is a useful statistic for a different purpose than pension funds would require. We do not consider other countries in this paper.

1.5 In section 2 we state the basic formula that we use. In section 3 we give estimation results. In section 4 we discuss further calculations and show the long term means and variances, and the correlation coefficients, comparing them with those of the original model. In section 5 we show certain results for forecasting. In section 6 we consider stochastic interpolation. In section 7 we show specimen forecasts. In section 8 we draw some conclusions. We put almost all the algebra into the Appendix.

## 2. Formulation

2.1 The original Wilkie model for retail prices, $Q(t)$, with parameters $QMU$, $QA$ and $QSD$, is;

$$I(t) = \ln(Q(t+1)) - \ln(Q(t))$$

$$IN(t) = I(t) - QMU$$

$$QE(t) = QSD.QZ(t)$$

$$I(t) = QA.IN(t) + QMU + QE(t) \quad (1)$$

$$Q(t+1) = Q(t).\exp(I(t))$$

2.2 The fuller form of the original model for wages, $W(t)$, with parameters $WMU$, $WW1$, $WW2$, $WA$ and $WSD$, is:

$$J(t) = \ln(W(t+1)) - \ln(W(t)$$

$$JN(t) = J(t) - WW1.I(t) - WW2.I(t-1) - WMU$$

$$WE(t) = WSD.WZ(t)$$

$$J(t) = WW1.I(t) + WW2.I(t-1) + WMU + WA.JN(\mathrm{t}) + WE(t) \quad (2)$$

$$W(t+1) = W(t).\exp(J(t))$$

Often $QA$ could be taken as zero, so the term $WA.JN(t)$ could be dropped out.

2.3 To include dependence of $I(t)$ on $J(t-1)$ we could elaborate the formula for $I(t)$ which we shall put in the form, changing the constant term from $QMU$ to $QM$:

$$I(t) = QA.I(t-1) + QW.J(t-1) + QM + QE(t) \quad (3)$$

But this suggests that a full symmetrical VAR model might be appropriate so we put:

$$J(t) = WQ.I(t-1) + WA.J(t-1) + WM + WE(t) \quad (4)$$

with a correlation coefficient, *QWR*, between *QZ*(*t*) and *WZ*(*t*). All the parameters for this can be estimated jointly. We assume throughout that *QZ*(t) and *WZ*(*t*) are distributed as unit normal, *N*(0, 1), although there is good evidence that the data is fatter tailed than normal.

2.4 It is convenient to replace *QMU* in formula (1) by *QM* in formula (3), since this allows us to estimate the parameters of (3) without reference to the *model* for *W*, though the *data* for *W*(*t*) is still used. We can then adjust formula (4) to:

$$J(t) = WQ.I(t-1) + WA.J(t-1) + WM + WBQ.QE(t) + WSD2.WZ(t) \quad (5)$$

Where the correlation is incorporated in the *WBQ* term and *WSD* is adjusted to *WSD*2. See Appendix A1.1. The parameters for formula (5) can then be estimated after formula (3) has been fitted.

2.5 Formula (3) and (4) or (3) and (5) are convenient for estimation, but for future simulation and other calculations it is convenient to re-arrange them as:

$$I(t) = QA.IN(t-1) + QW.JN(t-1) + QMU + QE(t) \quad (6)$$

$$J(t) = WQ.IN(t-1) + WA.JN(t-1) + WMU + WBQ.QE(t) + WSD2.WZ2(t) \quad (7)$$

with $IN(t) = I(t) - QMU$ and $JN(t) = J(t) - WMU$:

We can derive the values of *QMU* and *WMU* (see Appendix A2) from:

$$QMU = \{(1 - WA).QM + QW.WM\} / \{(1 - WA).(1 - QA) - QW.WQ\}$$

$$WMU = \{WQ.QM + (1 - QA).WM\} / \{(1 - WA).(1 - QA) - QW.WQ\}$$

## 3. Estimation

3.1 We have the annual June values of the Retail Prices Index (RPI) and its predecessors, and the Average Weekly Earnings Index (seasonally adjusted) from June 1921 to June 2016. We use these as *Q*(*t*) and *W*(*t*). This gives us annual values of *I*(*t*) and *J*(*t*) from 1922 to 2016, and we fit the model to the 94 values from 1923 to 2016, using the values for June 1922, in the formulae for 1923. We use maximum likelihood estimation (MLE), assuming normally distributed residuals. This gives the same estimates of the values of the parameters as least squares, but also gives us standard errors (and covariances) for all the parameters.

3.2 We start by estimating the parameters for formula (3), but omitting the term $QW.JN(t-1)$. This is the same as the present model, but *QM* has a different value from *QMU*. The results are shown in Table 1. We then include the omitted term, so that the results are now the same as with the joint, VAR, model. These results are also shown in Table 1. The improvement in log likelihood is significant at 3.01 and the extra term is also significantly different from zero. But the improvement is not enormous.

3.3 We show the skewness and kurtosis in Table 1. Both are large, for both models. This indicates that it is not correct to assume that the residuals are normally distributed, but we postpone investigating this until a later Part in this series.

3.4 We show at the foot of Table 1 the value of *QMU*, the long-term mean value of *I*(*t*). For the model omitting *QW*, *QMU* can be calculated as *QM* / (1 – *QA*). For the VAR model, including *QW*, it cannot be calculated until we also have the model for wages too, and then from the formula in section 2.5. The values of *QMU* are not the same, but are not very far apart.

Table 1. MLE parameter values and standard errors (s.e.) for models for *Q*.

| | Omitting *QW* | | Including *QW* | |
|---|---|---|---|---|
| Log likelihood | 259.05 | | 262.06 | |
| | MLE value | s.e. | MLE value | s.e. |
| *QA* | 0.571849 | 0.071778 | 0.273990 | 0.138151 |
| *QW* | | | 0.318039 | 0.127475 |
| *QM* | 0.018134 | 0.004785 | 0.013846 | 0.004943 |
| *QSD* | 0.038548 | 0.002812 | 0.037332 | 0.002723 |
| | | | | |
| Skewness | 1.30 | | 1.49 | |
| Kurtosis | 6.26 | | 6.78 | |
| | | | | |
| *QMU* | 0.042354 | | 0.043979 | |

3.5 We now do the same for the wages model, showing in Table 2 the results, first for the original model, and then for the new VAR model. This can be estimated either on its own, after the model for *Q* has been fitted, or jointly with the model for *Q*. The estimated values of the parameters are the same, except that for the joint model we obtain two different parameters, *WSD* and *QWR*, which are essentially the same as *WBQ* and *WSD*2 (see Appendix A1). The sum of the log likelihoods of the individual models is numerically the same as the (algebraically more complicated) log likelihood of the joint model. But the standard errors of the parameter estimates in the wages model differ, being larger in the joint model than in the individual one. This is presumably because in the joint model the values of the standard deviations, *QSD* and *WSD*, and the correlation coefficient, *QWR*, are estimated jointly, whereas in the individual wages model the values of *QSD* and the *QE*(*t*) terms are taken as fixed and only the parameter *WBQ* has to be estimated. In Table 2 we show the two standard errors as “s.e. (I)” for the individual model and “s.e. (VAR)” for the joint model. The standard errors in the model for *Q* are unaffected.

3.6 The skewness and kurtosis for the two models are also shown. These are not so large as for the model for *Q*, but still cast some doubt on the assumption of normally distributed residuals.

3.7 We show also at the foot of Table 2 the values of *WMU**, the long-term mean value of *J*(*t*). For the original model this is calculated as *WMU* + (*WW*1 + *WW*2).*QMU*. For the VAR model it is calculated from the formula in section 2.5 for *WMU*. (We call it *WMU** to avoid confusion with *WMU* in the original model that might be better called *WM*.)

3.8 The log likelihood for the two models is almost the same, that for the VAR model being 0.33 higher than for the original model, a trivial improvement. The same three connections are included. For the original model *I*(*t*) is included directly; for the VAR model it is connected through the correlation with *QZ*(*t*) or the *WBQ.QE*(*t*) term. *I*(*t*–1) is included directly in both models, and the estimates of the parameters, *WW*2 and *WQ* have quite similar values. *J*(*t*–1) is included directly in the VAR model, and indirectly through *WA.JN*(*t*–1) in the original model. The estimates of the parameters, called *WA* in both models, are rather different.

3.9 In aggregate the only improvement of the VAR model over the present one is in the model for *Q*. This is large enough to make it worth while using this model, at the expense of making the wages model essential, rather than an optional extra, and making the calculation of expected values and variances of future forecasts rather more complicated, as we see in section 4.

Table 2. MLE parameter values and standard errors (s.e.) for models for *W*.

| | Original model | | Formula (5) | | |
|---|---|---|---|---|---|
| Log likelihood | 312.47 | | 312.80 | | |
| | MLE value | s.e. | MLE value | s.e. (I) | s.e. (VAR) |
| *WW*1 | 0.567789 | 0.0597 | | | |
| *WW2* | 0.246563 | 0.0547 | | | |
| *WA* | 0.212206 | 0.0855 | | | |
| *WMU* | 0.020789 | 0.00393 | | | |
| *WSD* | 0.021837 | 0.00159 | | | |
| *WQ* | | | 0.262556 | 0.080519 | 0.111774 |
| *WA* | | | 0.380394 | 0.074298 | 0.103137 |
| *WM* | | | 0.023683 | 0.002880 | 0.003999 |
| *WBQ* | | | 0.561160 | 0.060115 | |
| *WSD*2 | | | 0.021759 | 0.001587 | |
| | | | | | |
| *WSD* | | | 0.030204 | | 0.002203 |
| *QWR* | | | 0.693584 | | 0.053525 |
| | | | | | |
| Skewness | 0.35 | | 0.25 | | |
| Kurtosis | 3.91 | | 4.21 | | |
| | | | | | |
| *WMU**, *WMU* | 0.055280 | | 0.056859 | | |

## 4. Forecast means and variances

4.1 For further calculations, for both the original models and the new VAR model, we round the parameter estimates shown in Tables 1 and 2 to four decimal places. We also adjust the values of the means *QMU* and *WMU*. We assume that *QMU* = 0.0250, as we did in Part 4 (Wilkie & Şahin, 2018), since that seems a more suitable value in current conditions, than the historic means of 0.0423 and 0.0440. For *WMU* we observe that *WMU* is 0.0130 greater than *QMU* for the new model and 0.0129 greater for the new model, so we choose a mean rate of wages growth of 0.0390 for both models. We put *WMU* = 0.0390 for the new model, but for the old model, we need an adjustment to give *WMU* = 0.01864. For the new model we also, for some purposes, use formula (5) for which we need *WBQ* and *WSD*2 instead of *WSD* and *QWR*, but to get consistent results we calculate values for these from the rounded values and use more decimal places.

4.2 Thus for the new model we use:

*QA* = 0.2740, *QW* = 0.3180, *WQ* = 0.2626, *WA* = 0.3804,
*QSD* = 0.0373, *WSD* = 0.0302, *QWR* = 0.6936
*QMU* = 0.0250, *WMU* = 0.0390
*WBQ* = 0.56157426, *WSD*2 = 0.02175489

For the old models we use:

*QA* = 0.5718, *QMU* = 0.0250, *QSD* = 0.0385
*WW*1 = 0.5678, *WW*2 = 0.2466, *WA* = 0.2122, *WMU* = 0.01864, *WSD* = 0.0218

4.3 In Part 2 we discussed the state variables, input and output variables, and initial conditions required for simulation. These remain almost the same, except that in the old model, the value of *I*(–1) was required in order to calculate *JN*(0), and it is no longer required for that purpose. However, it is required for other variables in the total model.

4.4 We wish to calculate the means, variances and covariances of future values of *I*(*t*), *QL*(*t*), *J*(*t*) and *WL*(*t*), conditional on the starting position at some time where we take *t* = 0. For numerical examples we use both neutral initial conditions and initial conditions as at June 2016. The initial conditions affect the means, but not the variances and covariances. This is all as shown in Part 2, but the formulae are more complicated for the VAR model. The algebra is discussed in Appendix A3. We also calculate the corresponding values for the present model. We show here only the results.

4.5 Neutral initial conditions for both old and new models are to put *I*(0) = *QMU* = 0.025, *W*(0) = (or *WMU**) = 0.390. The initial conditions as at June 2016 are:

*I*(0) = 0.0161, *I*(–1) = 0.0101, *Q*(0) = 263.1, *J*(0) = 0.0216, *W*(0) = 156.7.

We need *I*(–1) for the old model in order to calculate *JN*(0). For some purposes it is more convenient to use arbitrary values for *Q*(0) and *W*(0) such as 1.0, since these are multiplicative indices.

4.6 The long-term means and standard deviations of *I*(*t*) and *J*(*t*) are shown in Table 3: The means of each are the same for both models, because we have made them so. The standard deviations of each for the new model are a little lower than for the old, and in both cases the standard deviation for *J*(*t*) is lower than for *I*(*t*). The long-term correlation coefficient between *I*(*t*) and *J*(*t*) is very slightly lower in the new model. In total there is not very much difference.

Table 3. Long-terms means and standard deviations (SD) of *I*(*t*) and *J*(*t*), and correlation coefficient between *I*(*t*) and *J*(*t*)

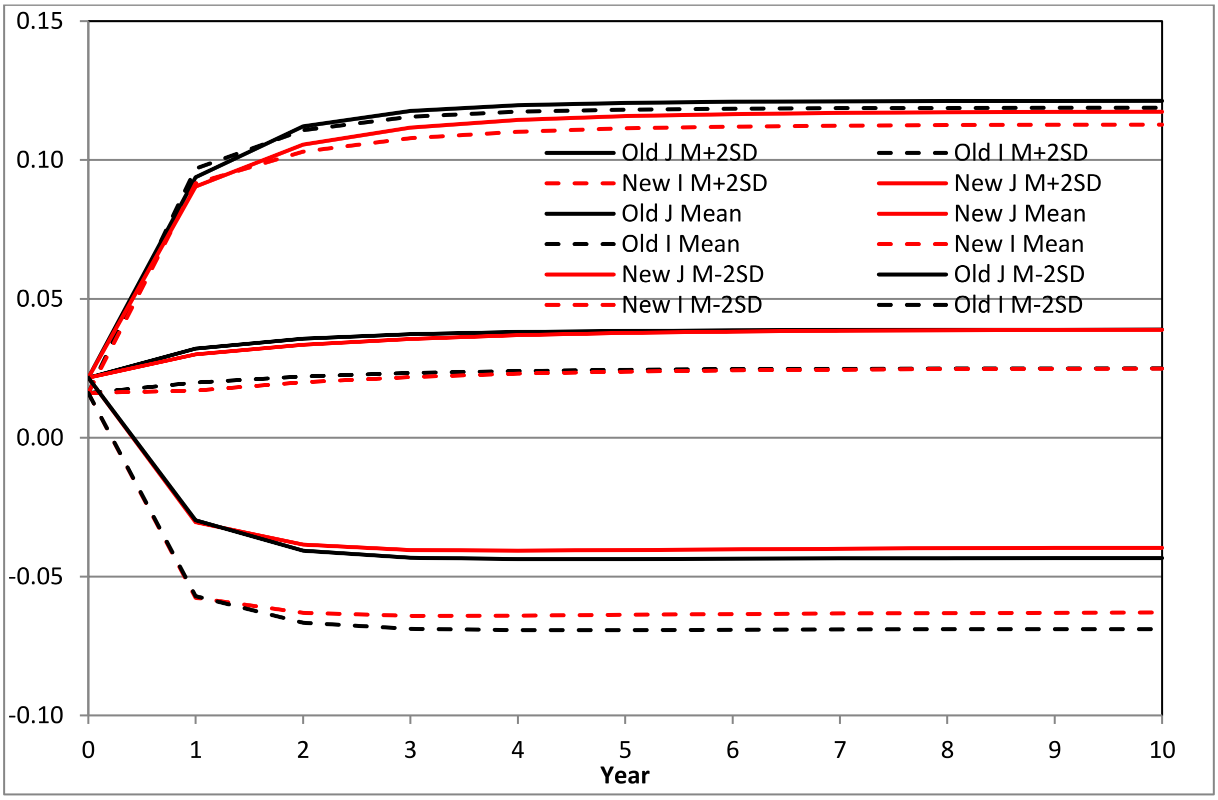
| | Old models | | | New VAR model | | |
|---|---|---|---|---|---|---|
| | Mean | SD | CC | Mean | SD | CC |
| *I* | 0.0250 | 0.0469 | 0.8081 | 0.0250 | 0.0439 | 0.7902 |
| *J* | 0.0390 | 0.0412 | | 0.0390 | 0.0392 | |

## 5. Forecasting

5.1 Using the formulae in the Appendix we calculate the means (formulae A1 to A4) and standard deviations (formulae A5 to A8 give the variances) of $I(t)$, $J(t)$, $QL(t)$ and $WL(t)$, for both the original model and the new VAR model We use the initial conditions as at June 2016. In Figure 1 we show the means and the means plus and minus twice the standard deviations of the forecast values of $I(t)$ and $J(t)$. If we assume normality of the innovations, this spread gives roughly a 95% confidence interval for the forecasts.

5.2 We go ahead only ten years, by which time all the values have converged almost to their long term values. We show the old model with black lines, the new with red, and show the values for $J$(t) as solid lines, and for $I(t)$ as dashed ones. The long-term means for both models are the same, because we made them so; our purpose is to show the spread. Consistent with the numbers in Table 3, the new model shows a slightly narrower spread than the old; more so for $I(t)$ than $J(t)$. Also the spread for $I(t)$ is greater than that for $J(t)$.

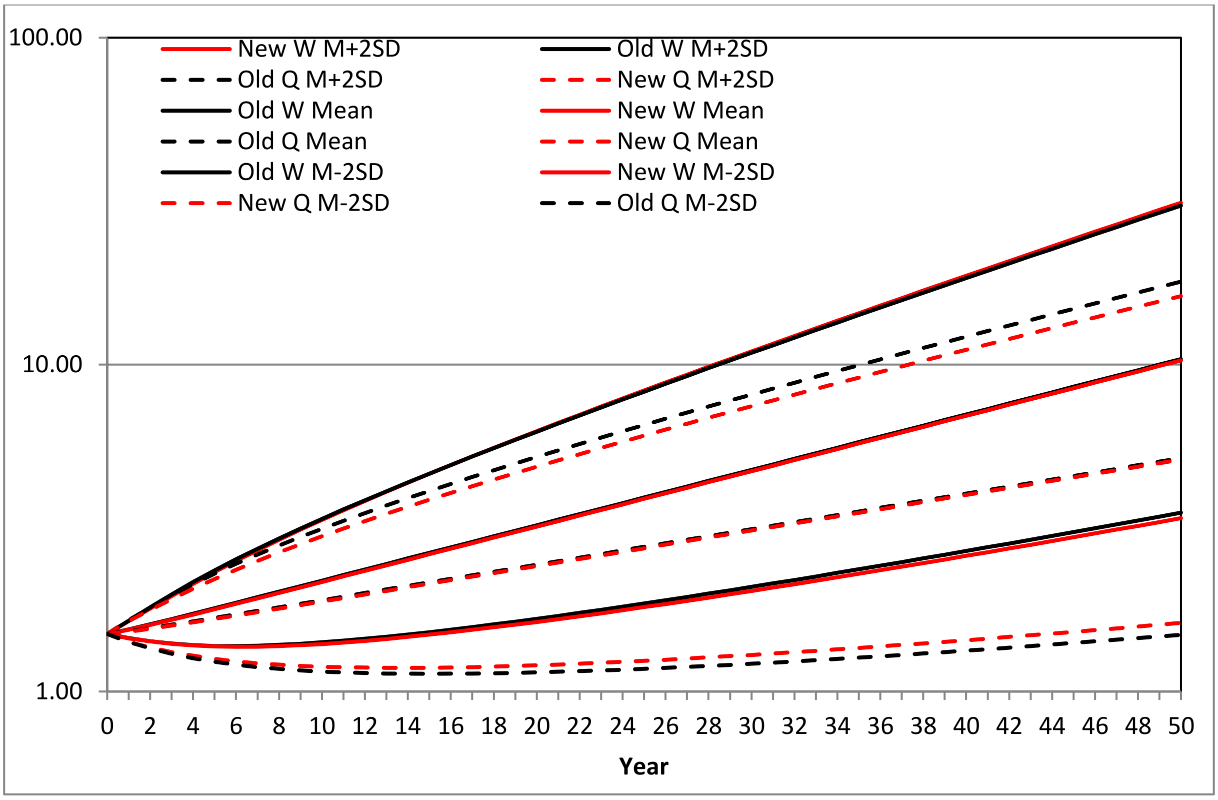
Figure 1. Mean and mean ± 2 × standard deviation, for $I(t)$ and $J(t)$ for both the old and the new models.




5.3 We do the same for $QL(t)$ and $WL(t)$, and then exponentiate to give the equivalent values of $Q(t)$ and $W(t)$. The result strictly does not give the mean values of $Q(t)$ and W(t), because, if we assume normality of the innovations, $Q(t)$ and $W(t)$ would be lognormally distributed, and what we call the mean is strictly the median; but the spread is still a 95% one. We show the results in Figure 2, for 50 years ahead. Since this is on a logarithmic vertical scale it shows in effect $QL(t)$ and $WL(t)$ on a linear scale. We choose arbitrary starring values for both $Q(0)$ and $W(0)$ of 1.5, in order to keep the results within two cycles of the scale.

5.4 The relative positions of the lines are the same as in Figure1. The new model shows a slightly narrower spread than the old, and the spread for *I*(*t*) is greater than that for *J*(*t*).

Figure 2. Mean and mean ± 2 × standard deviation, for *Q*(*t*) and *W*(*t*) for both the old and the new models.




## 6. Stochastic Interpolation

6.1 In Parts 3A, 3B and 3C we discussed stochastic interpolation, using Brownian bridges to obtain monthly values of *QL*(*t*,*m*) and *WL*(*t*,*m*), the simulated values in month *m* of year *t*. We update our analysis from June 2014 to June 2016, but the extra two years make little difference to the results. However, we observe that, in fitting the annual model, we start with *I*(1923), which is calculated from *QL*(1922) and *QL*(1923), so it is more consistent to use the monthly values starting from June 1922, giving one extra year at the beginning. For wages we explained in Part 3B why we could not start the monthly analysis till June 1934, and we do the same here. Thus we analyse *QL* for the 94 years from 1922 to 2016, and *WL* for the 82 years from 1934 to 2016.

6.2 For *QL* the monthly standard deviation for year *t* varies with the change in *QL*(*t*) over the year, *QLD*(*t*), which equals *I*(*t*+1), and the best formula is the same as before:

$$\sigma_m(t) = QSM(t) = QSA + QSB \times \text{Abs}(QLD(t) - QSC)$$

but with slightly altered parameters *QSA* = 0.004177, *QSB*= 0.055982 and *QSC* = 0.03933. The correlation coefficient is 0.5963, not very different from the previous 0.5865. The correlation coefficient between the standardised forward deviations for corresponding months in successive years is 0.5491, a small change from 0.5469 previously,

6.3 For *WL* we get a similar result, with rather little change from previously. The most suitable formula is still:

$$\sigma_m(t) = WSM(t) = WSA + WSB \times \text{Abs}(WLD(t))$$

with *WSA* = 0.002513 and *WSB*= 0.011482. The correlation coefficient is 0.3188, a small increase on the previous 0.2826. The correlation coefficient between the standardised forward deviations for corresponding months in successive years is 0.01933, a small change downwards from 0.2047 previously.

## 7. Simulations

7.1 We are now in apposition to simulate these variables for a chosen period ahead. In Figure 3 we show. First the actual values of *I*(*t*) and *J*(*t*), monthly, from January 1996 to June 2016. Note that *I*(*t*) and *J*(*t*) are the values of the changes in *QL*(*t*) and *WL*(*t*) over the preceding year, not over one month. Then we have done simulation of *I*(*t*) and *J*(*t*) on two bases, the old and the new, from which we calculate simulated values of *QL*(*t*) and *WL*(*t* ). We than apply stochastic interpolation over *QL*(*t*) and *WL*(*t*), using Brownian bridges as described in section 6; and from these we calculate values for *I*(*t*) and *J*(*t*) for intermediate months. We use the same random unit normal innovations for both the old and the new model.

7.2 We show the past actual data in black and red lines, and the simulated futures in other colours. We can see that the two models give quite similar future values. This is only one pair of simulations, and they may not be “typical” (if any single simulation can be thought to be typical). They show larger deviations of both variables, both upwards and downwards than the recent past data, but the upwards range is much less than that of the further back past, of the 1980s or the 1920s.

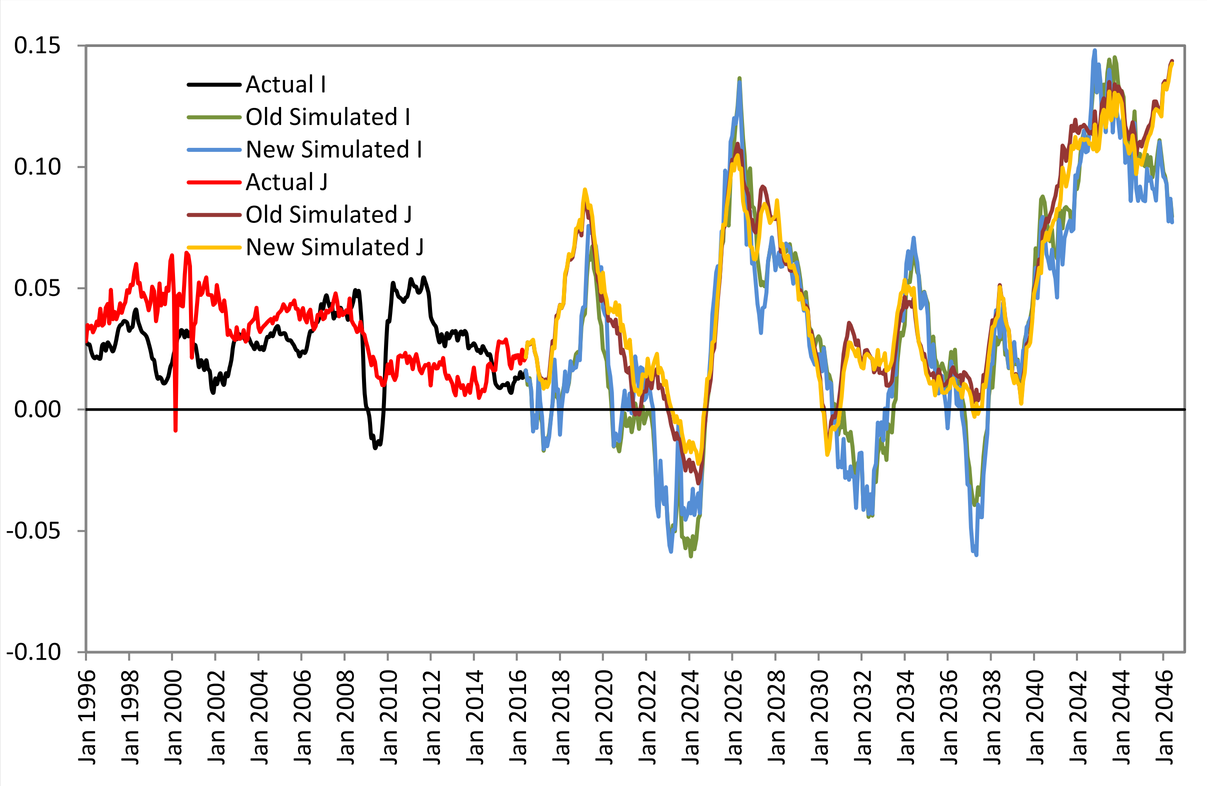
## 8. Conclusions

8.1 There are advantages and disadvantages of the new VAR model as compared with the old. The log likelihood for *I*(*t*) in the VAR model is rather higher, and the resulting standard deviation, *QSD*, rather smaller than for the old model, so it can be said to describe the data better. However, the change for *W*(*t*) is small. The VAR model is symmetrical, and is a standard time series model; but the old model is a normal transfer function model. Either model is straightforward to simulate, but one must note that the old “cascade-style” model can be simulated either variable by variable for a given number of years, or year by year for all variables. The VAR model requires one to proceed year by year.

8.2 With the VAR model, the formulae for forecast means and variances are more complicated than for the original model. We have not shown the formulae for all the other variables which depend on *I*(*t*), as we did in Part 2, but the complications spread everywhere.

8.3 The results for the new model are not very different from those using the old one, and one may well feel that a change in any existing system is not worth the trouble. But for someone starting a new development we would, on balance recommend the VAR one.

Figure 3. One pair of simulations, on both the old model and the new model, for *I*(*t*) and J(t) showing 20 years of past data and 30 years of simulated future values.




## References


Wilkie, A. D. (1986). A stochastic investment model for actuarial use. *T.F.A.*, **39**, 341-381.

Wilkie, A. D. (1995). More on a stochastic asset model for actuarial use. *B.A.J.*, **1**, 777-964.

Wilkie, A.D. Şahin Ş., Cairns, A.J.G. & Kleinow T. (2011). Yet more on a stochastic economic model: Part 1: Updating and Refitting, 1995 to 2009. *A.A.S.*, **5**, 53-99.

Wilkie, A.D. & Şahin Ş. (2016). Yet more on a stochastic economic model: Part 2: Initial Conditions, Select Periods and Neutralising Parameters. *A.A.S.*, 10, 1-51.

Wilkie, A.D. & Şahin Ş. (2017a). Yet more on a stochastic economic model: Part 3A Stochastic interpolation: Brownian and Ornstein-Uhlenbeck bridges. *A.A.S.*, **11, 74-99**.

Wilkie, A.D. & Şahin Ş. (2017b). Yet more on a stochastic economic model: Part 3B Stochastic bridging for retail prices and wages. *A.A.S.*, **11, 100-127**.

Wilkie, A.D. & Şahin Ş. (2017c). Yet more on a stochastic economic model: Part 3C Stochastic bridging for share yields and dividends and interest rates. *A.A.S.*, **11, 128-163**.

Wilkie, A.D. & Şahin Ş. (2018). Yet more on a stochastic economic model: Part 4 A model for share earnings, dividends and prices. *A.A.S.*, **12**, xx-xx.

**Appendix**

A1 *Equivalence of formulae (4) and (5)*

A1.1 In section 2.4 we state that formula (5) for $J(t)$ is equivalent to formula (4). The formulae are the same except for the ending terms. Formula (4) ends with $WE(t)$, which equals $WSD.WZ(t)$, and there is correlation of $QWR$ between $WZ(t)$ and $QZ(t)$. Formula (5) ends with $WBQ,QE(t)$ + $WSD2.WZ2(t)$, and $WZ2(t)$ is assumed to be independent of $QZ(t)$. The equivalence is easily shown if we consider the Choleski decomposition of the correlation matrix:

$$M = \begin{bmatrix} 1 & CQW \\ CQW & 1 \end{bmatrix}$$

which is

$$C = \begin{bmatrix} 1 & 0 \\ CQW & \sqrt{1 - CQW^2} \end{bmatrix}$$

and $C.C^T = M$, so that $C$ is, in a sense, the square root of $M$.

We now put $WBQ = QWR.WSD \;/\; QSD$ and $WSD2 = WSD.\sqrt{1 - CQW^2}$.

A 1.2 For both formulae the terms we are considering are the only stochastic ones, and in both cases the expected value is zero. The formula (4) term is $WSD.WZ(t)$ and the variance of this is $WSD^2$. The corresponding stochastic part of $I(t)$, in formula (3) is $QSD.QZ$(t). The covariance of $I(t)$ and $J(t)$ is therefore $QSD.WSD.\text{Cov}[QZ(t).WZ(t)] = QSD.WSD.QWR$.

A1.3 We expand formula (5) to give:

$$WBQ.QSD.QZ(t) \text{ `}+ WSD2.WZ2(t) = QWR.WSD.QZ(t) + WSD.\sqrt{1 - CQW^2}.\ WZ2(t)$$

The variance of this is:

$$(WBQ.QSD)^2 + WSD2^2 = (QWR.\text{WSD})^2 + \text{WSD}^2.(1 - QWR^2)$$

$$= WSD^2.\ (QWR^2 + 1 - QWR^2)$$

$$= WSD^2$$

which is the same as the variance in formula (4). The covariance between $J(t)$ with formula (5) and $I(t)$ is $QWR.WSD.QSD$, which is also the same as in formula (4).

A1.4 In our estimation for the period 1923 to 2016 with the VAR model we get $QSD$ = 0.037322, $WSD$ = 0.030204 and $QWR$ = 0.693584 (see Tables 1 and 2). We can calculate from the formulae above that $WBQ = QWR.WSD \;/\; QSD$ = 0.561160 and also that $WSD2 = WSD.\sqrt{1 - CQW^2}$ = 0.021759, which agree exactly with those parameters estimated with the individual model. Using our rounded values for the parameters we get $WBQ$ = 0.056157426, and $WSD2$ = 0.02175489.

A2 *Calculation of* QMU *and* WMU

A2.1 In section 2.5 we state that formula (3) and (5) can be transformed to formulae (6) and (7). The formulae are:

(3) $I(t) = QA.I(t-1) + QW.J(t-1) + QM + QE(t)$

(5) $J(t) = WQ.I(t-1) + WA.J(t-1) + WM + WBQ.QE(t) + WSD2.WZ(t)$

(6) $I(t) = QA.IN(t-1) + QW.JN(t-1) + QMU + QE(t)$

(7) $J(t) = WQ.IN(t-1) + WA.JN(t-1) + WMU + WBQ.QE(t) + WSD2.WZ2(t)$

with $IN(t) = I(t) - QMU$ and $JN(t) = J(t) - WMU$.

A2.2 Equating corresponding formulae we see that:

$$QM = -QA.QMU - QW.WMU + QMU = (1 - QA).QMU - QW.WMU$$

$$WM = -WQ.QMU - WA.WMU + WMU = -WQ.QMU + (1 - WA).WMU$$

From these simultaneous equations we get:

$$QMU = \{(1 - WA).QM + QW.WM\} / \{(1 - WA).(1 - QA) - QW.WQ\}$$

$$WMU = \{WQ.QM + (1 - QA).WM\} / \{(1 - WA).(1 - QA) - QW.WQ\}$$

A2.3 For our data we have $QA = 0.273990$, $QW = 0.318039$, $WQ = 0.262556$, $WA = 0.380294$, $QM = 0.013846$ and $WM = 0.023683$, so we get $QMU = 0.043979$ and $WMU = 0.056859$.

A3 *Means of forecasts*

A3.1 Our methods here are similar to the well-known method of solving for a univariate recurrence relation like:

$$X_t = a.X_{t-1} + b.X_{t-2}$$

where we find the roots, $x_1$ and $x_2$, of the characteristic equation:

$$x^2 = a.x + b$$

We then put $X_t = c_1.x_1^t + c_2.x_2^t$

And find $c_1$ and $c_2$ from any two given values of $X_t$.

A3.2 If we ignore the stochastic parts of the equations for $I(t)$ and $J(t)$ and omit the mean terms we have a bivariate recurrence relation for future expected values;

$IN(t) = QA.IN(t-1) + QW.JN(t-1)$

$JN(t) = WQ.IN(t-1) + WA.JN(t-1)$

We can write this in matrix form as:

$\boldsymbol{xN}(t) = \mathbf{T}.\boldsymbol{xN}(t{-}1)$

where $\boldsymbol{xN}(t) = [IN(t), JN(t)]^T$

$$\boldsymbol{T} = \begin{bmatrix} QA & QW \\ WQ & WA \end{bmatrix}$$

A3.3 The equivalent of the roots, $x_1$ and $x_2$ are the eigenvalues, $\lambda_1$ and $\lambda_2$, of this matrix, obtained from $|\boldsymbol{T} - \lambda.\boldsymbol{I}| = 0$

$$\begin{vmatrix} QA - \lambda & QW \\ WQ & WA - \lambda \end{vmatrix} = 0$$

$(QA - \lambda).(WA - \lambda) - QW.WQ = 0$

$\lambda_1, \lambda_2 = = \{(QA + WA) \pm \sqrt{((QA - WA)^2 + 4.QW.WQ)}\}/2$

A3.4 For our data we get $\lambda_1$, $\lambda_2$ = 0.327192 ± 0293854 so $\lambda_1$ = 0.654383 and $\lambda_2$ = 0.033366, these are both positive and real, which is convenient. But these are not the only possibilities, as we discuss further in A3.8

A3.4 We assume that we start at time $t = 0$ with initial conditions $I(0)$ and $J(0)$ known. We can then calculate $IN(0) = I(0) - QMU$ and $JN(0) = J(0) - WMU$. Using formulae (6) and (7) we can calculate $E[IN(1)] = QA.IN(0) + QW.JN(0)$ and $E[JN(1)] = WQ.IN(0) + WA.JN(0)$ and recursively $E[IN(t+1)] = QA.IN(t) + QW.JN(t)$ and $E[JN(t+1)] = WQ.IN(t) + WA.JN(t)$. At this stage we omit all stochastic terms. We have a simple bivariate recurrence relation for $E[IN(t)]$ and $E[JN(t)]$

A3.5 We now postulate that $E[IN(t)] = A1.\lambda_1^t + A2.\lambda_2$ and $E[JN(t)] = B1.\lambda_1^t + B2.\lambda_2^t$. We know $E[IN(0)] = IN(0)$ and we calculate $E[IN(1)]$ as above. We can then put:

$E[IN(0)] = A1 + A2$

and

$E[IN(1)] = QA.IN(0) + QW.JN(0) = A1.\lambda_1 + A2.\lambda_2$

From these we readily get:

$A1 = \{(QA - \lambda_2).IN0 + QW.JN(0)\} / (\lambda_1 - \lambda_2)$

$A2 = IN(0) - A1 = \{(QA - \lambda_1).IN0 + QW.JN(0)\} / (\lambda_2 - \lambda_1)$

Then similarly we get:

$B1 = \{(WA - \lambda_2).JN0 + WQ.IN(0)\} / (\lambda_1 - \lambda_2)$

$B2 = JN(0) - B1 = \{(WA - \lambda_1).JN0 + WQ.IN(0)\} / (\lambda_2 - \lambda_1)$

A3.6 The proof by induction that, if $E[IN(t)] = A1.\lambda_1^t + A2.\lambda_2^t$ then $E[IN(t+1)] = A1.\lambda_1^{t+1} + A2.\lambda_2^{t+1}$ is tedious, and relies on the fact that $\lambda_1$ and $\lambda_2$ are the roots of the quadratic in $\lambda$, so that $\lambda_1 + \lambda_2 = QA + WA$ and $\lambda_1.\lambda_2 = QA.WA - QW.WQ$. There is a similar formula for $E[JN(t)]$.

A3.7 We can now, given the relevant values of $IN(0)$ and $JN(0)$, readily calculate $E[IN(t)]$ and $E[JN(t)]$ for any $t$, and hence also $E[I(t)]$ and $E[J(t)]$, which are:

$$E[I(t)] = A1.\lambda_1^t + A2.\lambda_2^t + QMU \quad \text{(A1)}$$

$$E[J(t)] = B1.\lambda_1^t + B2.\lambda_2^t + WMU \quad \text{(A2)}$$

We see also that as $t \to \infty$, $E[IN(t)]$ and $E[JN(t)] \to 0$, hence $E[I(t)] \to QMU$ and $E[J(t)] \to WMU$. These results depend on $|\lambda_1| < 1$ and $|\lambda_2| < 1$, which is true in our case.

A3.8 In A3.4 we noted that $\lambda_1$ and $\lambda_2$ were both are positive and real, but that there were other possibilities with different data. Since, with different data, we would expect to have different parameter estimates, then we might have one of the roots greater than unity, so the system would be divergent rather than mean-reverting. If the square root part were a little larger, we might get $\lambda_2$ negative, which is not much trouble, since the formulae are the same, except that the effect of $\lambda_2^t$ alternates between positive and negative. But if $\lambda_2$ remains small, then the terms in $\lambda_2^t$ reduce very rapidly, and the dominant effect is that of $\lambda_1$. If $\lambda_2$ were exactly 0, then the transition matrix would be singular, and there would be a problem in identifying the parameters uniquely. Another possibility is that the term inside the square root is negative so the roots are a complex conjugate pair. We have not explored this fully, but the algebra would be the same, with complex numbers throughout, and the effect would probably be that the forecast values would lie on a damped sine wave.

A3.9 The forecast means for $I(t)$ and $J(t)$ have formulae analogous to those for $I(t)$ in the original model, as shown in Part 2, except that we have terms with powers of $\lambda_1$ and $\lambda_2$ instead of terms in $QA$ In Appendix A2.2 of that paper we had:

$$E[QL(t)] = QL(0) + t.QMU + QA.\text{SumGS}(QA,t).IN(0)$$

where $\text{SumGS}(x, t) = 1 + x + x^2 + \ldots + x^{t-1} = (1 - x^t) / (1 - x)$

We can then put, in our case:

$$E[QL(t)] = QL(0) + t.QMU + A1.\lambda_1.\text{SumGS}(\lambda_1,t) + A2.\lambda_2.\text{SumGS}(\lambda_2,t) \quad \text{(A3)}$$

$$E[WL(t)] = WL(0) + t.WMU + B1.\lambda_1.\text{SumGS}(\lambda_1,t) + B2.\lambda_2.\text{SumGS}(\lambda_2,t) \quad \text{(A4)}$$

As $t \to \infty$, $E[QL(t)] \to \infty$ unless $QMU = 0$ and $E[WL(t)] \to \infty$ unless $WMU = 0$.

If $QMU = 0$ then:

$$E[QL(t)] \to QL(0) + A1.\lambda_1 / (1 - \lambda_1) + A2.\lambda_2 / (1 - \lambda_2)$$

If $WMU = 0$ then:

$$E[WL(t)] \to WL(0) + B1.\lambda_1 / (1 - \lambda_1) + B2.\lambda_2 / (1 - \lambda_2)$$

A4 *Psi-weights and variances of forecasts*

A4.1 We now wish to calculate the variances, and later the covariances, of the forecast variables. For this we can generally ignore the deterministic parts that we have used in the forecast means. As we stated in part 2, we can put:

$$I(t) = E[I(t)] + \{\textstyle\sum_{i=1,t} \Psi_{IQ}(i).QSD.QZ(t+1-i)\} + \{\textstyle\sum_{i=1,t} \Psi_{IW}(i).WSD.WZ(t+1-i)\}$$

whence we derive:

$$\begin{aligned} \mathrm{Var}[I(t)] = \mathrm{Var}[IN(t)] = {} & \{\textstyle\sum_{i=1,t} \Psi_{IQ}(i)^2\}.QSD^2 + \{\textstyle\sum_{i=1,t} \Psi_{IW}(i)^2\}.WSD^2 \\ & + \{\textstyle\sum_{i=1,t} \Psi_{IQ}(i)\, \Psi_{IW}(i)\}.QWR.QSD.WSD \end{aligned}$$

where the last term recognises the correlation of $QWR$ between $QE(t)$ and $WE(t)$, with a covariance of $QWR.QSD.WSD$.

We have similar formulae for $J(t)$ and for $\mathrm{Var}[J(t)]$.

A4.2 We use the formulae:

$$I(t) = QA.IN(t-1) + QW.JN(t-1) + QMU + QSD.QZ(t)$$

$$J(t) = WQ.IN(t-1) + WA.JN(t-1) + WMU + WSD.WZ(t)$$

with the covariance $E[QZ(t), WZ(t)] = QWR$.

A 4.3 We see immediately that, for $t = 1$:

$\Psi_{IQ}(1) = 1$
$\Psi_{IW}(1) = 0$
$\Psi_{JQ}(1) = 0$
$\Psi_{JW}(1) = 1$

and then for $t = 2$ that:

$\Psi_{IQ}(2) = QA$
$\Psi_{IW}(2) = QW$
$\Psi_{JQ}(2) = WQ$
$\Psi_{JW}(2) = WA$

but we find that these formulae very quickly get complicated.

A4.4 We now postulate that:

$\Psi_{IQ}(t) = C1.\lambda_1^{t-1} + C2.\lambda_2^{t-1}$
$\Psi_{IW}(t) = D1.\lambda_1^{t-1} + D2.\lambda_2^{t-1}$
$\Psi_{JQ}(t) = E1.\lambda_1^{t-1} + E2.\lambda_2^{t-1}$
$\Psi_{JW}(t) = F1.\lambda_1^{t-1} + F2.\lambda_2^{t-1}$

Note that the power applied to of $\lambda_1$ and $\lambda_2$ is one less than for E[*IN*(*t*)] and E[*JN*(*t*)]

Using the $\Psi$ values for $t = 1$ and $t = 2$ we get:

$C1 = (QA\text{–}\ \lambda_2/\ (\lambda_1 - \lambda_2)$
$C2 = 1 - C1 = 1 - (QA\text{–}\ \lambda_2) / (\lambda_1 - \lambda_2) = (\lambda_1 - QA) / (\lambda_1 - \lambda_2)$
$D1 = QW / (\lambda_1 - \lambda_2)$
$D2 = -D1 = -QW / (\lambda_1 - \lambda_2)$
$E1 = WQ / (\lambda_1 - \lambda_2)$
$E2 = -E1 = -WQ / (\lambda_1 - \lambda_2)$
$F1 = (WA\text{–}\ \lambda_1) / (\lambda_1 - \lambda_2)$
$F2 = 1 - F1 = 1 - (WA\text{–}\ \lambda_2)$

We can prove the formulae for general *t* by induction, rather laboriously.

A4.5 To get the $\Psi$ weights for *QL* and *WL* we use the same methods as in Part 2 for *QL*, getting:

$\Psi_{QLQ}(t) = C1.\text{SumGS}(\lambda_{1,}t) + C2.\ \text{SumGS}(\lambda_2,t)$
$\Psi_{QLW}(t) = D1.\text{SumGS}(\lambda_{1,}t) + D2.\ \text{SumGS}(\lambda_2,t)$
$\Psi_{WLQ}(t) = E1.\text{SumGS}(\lambda_{1,}t) + E2.\ \text{SumGS}(\lambda_{2,}t)$
$\Psi_{JWLW}(t) = F1.\text{SumGS}(\lambda_{1,}t) + F2.\text{SumGS}(\lambda_{2,}t)$

A4.6 We can now calculate the variances, again following the methods for these in Part 2, noting that, instead of terms with powers only of $QA^2$ multiplying $QSD^2$, we have powers of $\lambda_1^2$, $\lambda_2^2$ and $\lambda_1.\lambda_2$, multiplying terms in $QSD^2$, $WSD^2$, and the covariance $QWR.QSD.WSD = CV$

$$\begin{aligned}\text{Var}[I(t)] = \{\textstyle\sum_{i=1,t} \Psi_{IQ}(i)^2\}.QSD^2 + \{\textstyle\sum_{i=1,t} \Psi_{IW}(i)^2\}.WSD^2 \\ + 2\{\textstyle\sum_{i=1,t} \Psi_{IQ}(i).\ \Psi_{IW}(i)^2\}.CV\end{aligned} \qquad \text{(A5)}$$

$$\begin{aligned}= \{&C1^2.\text{SumGS}(\lambda_1^2,t) + C2^2.\text{SumGS}(\lambda_2^2,t)+ 2C1.C2.\text{SumGS}(\lambda_1.\lambda_2,t)\}.QSD^2 \\ &+ \{D1^2.\text{SumGS}(\lambda_1^2,t) + D2^2.\text{SumGS}(\lambda_2^2,t) + 2D1.D2.\text{SumGS}(\lambda_1.\lambda_2,t)\}.WSD^2 \\ &+ 2\{C1.D1.\text{SumGS}(\lambda_1^2,t) + C2.D2.\text{SumGS}(\lambda_2^2,t) \\ &\quad + (C1.D2 + C2.D1).\text{SumGS}(\lambda_1.\lambda_2,t)\}.CV\end{aligned}$$

$$\begin{aligned}= \{&C1^2.QSD^2 + D1^2.WSD^2 + 2C1.D1.CV\}.\text{SumGS}(\lambda_1^2,t) \\ &+ \{C2^2.QSD^2 + D2^2.WSD^2 + 2C2.D2.CV\}.\text{SumGS}(\lambda_2^2,t) \\ &+ 2\{C1.C2.QSD^2 + D1.D2.WSD^2 + (C1.D2 + C2.D1).CV\}.\text{SumGS}(\lambda_1.\lambda_2,t)\end{aligned}$$

$$= IVA.\text{SumGS}(\lambda_1^2,t) + IVB.\text{SumGS}(\lambda_2^2,t) + IVC.\text{SumGS}(\lambda_1.\lambda_2,t)$$

where

$IVA = C1^2.QSD^2 + D1^2.WSD^2 + 2C1.D1.CV$
$IVB = C2^2.QSD^2 + D2^2.WSD^2 + 2C2.D2.CV$
$IVC = 2\{C1.C2.QSD^2 + D1.D2.WSD^2 + (C1.D2 + C2.D1).CV\}$

$$\text{Var}[J(t)] = JVA.\text{SumGS}(\lambda_1^2,t) + JVB.\text{SumGS}(\lambda_2^2,t) + JVC.\text{SumGS}(\lambda_1.\lambda_2,t) \quad \text{(A6)}$$

where

$JVA = E1^2.QSD^2 + F1^2.WSD^2 + 2E1.F1.CV$
$JVB = E2^2.QSD^2 + F2^2.WSD^2 + 2E2.F2.CV$
$JVC = 2\{E1.E2.QSD^2 + F1.F2.WSD^2 + (E1.F2 + F2.E1).CV\}$

As $t \to \infty$ $\quad \text{Var}[I(t)] \to IVA / (1 - \lambda_1^2) + IVB / (1 - \lambda_2^2) + IVC / (1 - \lambda_1.\lambda_2)$

$\text{Var}[J(t)] \to JVA / (1 - \lambda_1^2) + JVB / (1 - \lambda_2^2) + JVC / (1 - \lambda_1.\lambda_2)$

$$\begin{aligned}\text{Var}[QL(t)] = {} & IVA.\{t - 2\lambda_1\text{SumGS}(\lambda_1,t) + \lambda_1^2\text{SumGS}(\lambda_1^2,t)\} / (1-\lambda_1)^2 \\ & + IVB.\{t - 2\lambda_2\text{SumGS}(\lambda_2,t) + \lambda_2^2\text{SumGS}(\lambda_2^2,t)\} / (1-\lambda_2)^2 \\ & + IVC.\{t - \lambda_1\text{SumGS}(\lambda_1,t) - \lambda_2\text{SumGS}(\lambda_1,t) + \lambda_1.\lambda_2\text{SumGS}(\lambda_1.\lambda_2,t)\} / (1-\lambda_1)(1-\lambda_2) \quad \text{(A7)}\end{aligned}$$

$$\begin{aligned}\text{Var}[WL(t)] = {} & JVA.\{t - 2\lambda_1\text{SumGS}(\lambda_1,t) + \lambda_1^2\text{SumGS}(\lambda_1^2,t)\} / (1-\lambda_1)^2 \\ & + JVB.\{t - 2\lambda_2\text{SumGS}(\lambda_2,t) + \lambda_2^2\text{SumGS}(\lambda_2^2,t)\} / (1-\lambda_2)^2 \\ & + JVC.\{t - \lambda_1\text{SumGS}(\lambda_1,t) - \lambda_2\text{SumGS}(\lambda_1,t) + \lambda_1.\lambda_2\text{SumGS}(\lambda_1.\lambda_2,t)\} / (1-\lambda_1)(1-\lambda_2) \quad \text{(A8)}\end{aligned}$$

As $t \to \infty$ $\quad \text{Var}[QL(t)] \to \infty$ and $\text{Var}[WL(t)] \to \infty$

## A5 *Covariances of forecasts*

5.1 We start with:

$$\begin{aligned}\text{Covar}[I(t),J(t)] = {} & \textstyle\sum_{i=1,t} \{\Psi_{IQ}(i).\Psi_{JQ}(i)\}.QSD^2 + \sum_{i=1,t} \{\Psi_{IW}(i).\Psi_{Jw}(i)\}.WSD^2 \\ & + \textstyle\sum_{i=1,t} \{\Psi_{IQ}(i).\Psi_{JW}(i) + \Psi_{IW}(i).\Psi_{JQ}(i)\}.CV\end{aligned}$$

and after a great deal of manipulation we get:

$$\begin{aligned}\text{Covar}[I(t),J(t)] = {} & QWVA.\text{SumGS}(\lambda_1^2,t) + QWVB.\text{SumGS}(\lambda_2^2,t) \\ & + QWVC\ \text{SumGS}(\lambda_1.\lambda_2,t)\end{aligned}$$

where

$QWVA = C1.E1.QSD^2 + D1.F1.WSD^2 + (C1.F1 + D1.E1).CV$
$QWVB = C2.E2.QSD^2 + D2.F2.WSD^2 + (C2.F2 + D2.E2).CV$
$QWVC = (C2.E1 + C1.E2).QSD^2 + (D2.F1 + D1.F2).WSD^2$
$\quad + (C2.F1. + C1.F2 + D2\ E1 + D1.E2).CV$

As $t \to \infty$ $\quad \text{Covar}[I(t), J(t)] \to QWVA / (1 - \lambda_1^2) + QWVB / (1 - \lambda_2^2) + QWVC / (1 - \lambda_1.\lambda_2)$

Similarly

$$\text{Covar}[QL(t),WL(t)] = QWVA.\ \{t - 2\lambda_1\text{SumGS}(\lambda_1,t) + \lambda_1^2\text{SumGS}(\lambda_1^2,t)\} / (1-\lambda_1)^2$$
$$+ QWVB.\ \{t - 2\lambda_2\text{SumGS}(\lambda_2,t) + \lambda_2^2\text{SumGS}(\lambda_2^2,t)\} / (1-\lambda_2)^2$$
$$+ QWVC.\ \{t - \lambda_1\text{SumGS}(\lambda_1,t) - \lambda_2\text{SumGS}(\lambda_1,t) + \lambda_1.\lambda_2\text{SumGS}(\lambda_1.\lambda_2,t)\} / (1-\lambda_1)(1-\lambda_2)$$

As $t \to \infty$ Covar$[I(t), J(t)] \to \infty$

A6 *Final note*

6.1 Many of the results in this Appendix can be expressed in matrix notation, and if the number of variables were large this would be more compact. However, for practical calculation it seems necessary to work through the eigenvalues anyway, and since these are easily calculated for a 2 by 2 matrix, as is the inverse, we consider it neater, in this case, to avoid matrix notation in most places.

END