TI 2012-025/4
Tinbergen Institute Discussion Paper

Evidence on Features of a DSGE Business Cycle Model from Bayesian Model Averaging

Rodney Strachan[1]
Herman K. van Dijk[2]

[1] Research School of Economics, The Australian National University;
[2] Econometric and Tinbergen Institutes, Erasmus University Rotterdam, and VU University Amsterdam.

Evidence on Features of a DSGE Business Cycle model from Bayesian Model Averaging[1,2]

Rodney W. Strachan

Research School of Economics, The Australian National University

Herman K. van Dijk

Econometric and Tinbergen Institutes, Erasmus University Rotterdam and VU University Amsterdam

ABSTRACT

The empirical support for features of a Dynamic Stochastic General Equilibrium model with two technology shocks is evaluated using Bayesian model averaging over vector autoregressions. The model features include equilibria, restrictions on long-run

[1] *Manuscript received October 2008; revised September 2010.

[2] We would like to thank Frank Schorfheide and three anonymous referees for extensive comments which led to a substantial revision and extension of the results presented in the original working paper. A preliminary version of this paper has been presented at the US Federal Reserve Bank in 2008 and at Cambridge University in 2011. We also thank Luc Bauwens, Geert Dhaene, John Geweke, David Hendry, Lennart Hoogerheide, Soren Johansen, Helmut Lutkepohl, Christopher Sims, Mattias Villani and Anders Warne for helpful discussions on the topic of this paper. Of course, any remaining errors remain the responsibility of the authors. Van Dijk acknowledges financial support from the Netherlands Organization of Scientific Research.

responses, a structural break of unknown date and a range of lags and deterministic processes. We find support for a number of features implied by the economic model and the evidence suggests a break in the entire model structure around 1984 after which technology shocks appear to account for all stochastic trends. Business cycle volatility seems more due to investment specific technology shocks than neutral technology shocks.

Key Words: Posterior probability; Dynamic stochastic general equilibrium model; Cointegration; Model averaging; Stochastic trend; Impulse response; Vector autoregressive model.

JEL Codes: C11, C32, C52

Running Head: Evidence on a DSGE model from BMA

1 Introduction.

In this paper we evaluate the robustness, in face of model uncertainty, of the empirical support for long-run equilibrium relationships implied by a Dynamic Stochastic General Equilibrium (DSGE) business cycle model subject to a investment-specific technology shock and a neutral technology shock. We embed these features of the DSGE model within a set of Vector AutoRegressive (VAR) models and take into account model uncertainty using a Bayesian Model Averaging (BMA) approach. Our work is distinguished from most other model averaging papers since averaging over systems of variables (rather that single equation models) implies averaging over features of the model rather than averaging over sets of regressors. Averaging over model features in systems of variables adds a level of complexity compared to single equation models and requires careful consideration of prior distributions. However, the approach we propose makes such an exercise feasible and the empirical results suggest the exercise is worthwhile.

The DSGE model investigated in this paper is based upon one described by Fisher (2006). This model has several features that can be empirically weighted: for example, the economic model suggests that the Great Ratios (e.g., consumption to income, investment to income) are stationary, that only investment-specific technology shocks have permanent effects on the real investment good price, and only technology shocks affect productivity in the long run. Model uncertainty derives from uncertainty over the number of stochastic trends, on the form of the reduced form equilibrium (cointegrating) relations, on the form of the deterministic trends, and, finally, on lag length. By considering the unconditional evidence, where 'unconditional' means that the empirical evidence does not depend upon a single model, it is possible to identify those features that have stronger empirical support. The joint evidence for those features implied by the model will indicate its empirical support.[3]

The idea underlying BMA is relatively straightforward. Model specific estimates are weighted by the corresponding posterior model probability and then averaged over the set of models considered. Although many statistical arguments have been made in the literature to support model averaging (e.g., Leamer 1978, Hodges 1987, Draper 1995, Min and Zellner 1993 and Raftery, Madigan and Hoeting 1997), an increasing number of recent applications suggest its relevance for macroeconometrics (Fernández, Ley and Steel 2001, Sala-i-Martin, Doppelhoffer and Miller 2004, Koop and Potter 2003, Wright 2008 and Koop, León-González and Strachan 2012). There are several arguments for model averaging and only a few are mentioned here. At the simplest level, it is often attractive to report inferences robust to model specification. Second, a large body of applied work has demonstrated that averaging results in gains in forecasting accuracy (Bates and Granger 1969, Diebold and Lopez 1996,

[3]We use the word 'features' rather than 'structures' to avoid confusing our work with structural VAR analysis, although we later consider 'structural breaks' in the common use sense of this term.

Newbold and Harvey 2001, Terui and van Dijk 2002, Hoogerheide, Kleijn, Ravazzolo, van Dijk and Verbeek 2010 and Wright 2008). Some explanation for this phenomenon in particular cases was provided by Hendry and Clements (2002). Thirdly, from a methodological point of view, averaging over models addresses to some degree the well understood pre-test problem (see, for example, Poirier 1995, pp. 519-523).

There is clear evidence from the literature that a structural break should be considered around 1984 (see, for example, McConnell and Perez-Quiros 2000 and Stock and Watson 2002). This literature suggests there is evidence of a break in possibly both the variances and mean equation coefficients. There is little work to date on changes in the overall structure or features of the model, such as changes in lag dynamics or stability of variables. We find the empirical evidence suggests that allowing for structural changes in the models, that is allowing the process to switch from one model to another, rather than just the parameter values, is justified.[4] Incorporating this model switching is a computationally challenging task. We simplify this task by considering what was the most likely model before the break and which was the most likely model after the break, rather than trying to compute the evidence of the switch from one model structure to another.

This paper makes three contributions. First, we show how to obtain posterior

[4] The introduction of the structural break analysis was on the suggestion of a referee to whom we are very grateful. This extension has significantly altered the results and conclusions.

inference from model averages in which the economically and econometrically important features may have weights other than zero or one. In other words, the inferences are based on a large finite mixture of model structures. Second, this paper treats a structural break as a change in the entire structure of the model, not just a change in parameter values. We find strong evidence that the entire structure, rather than just the parameter values, has changed. This extension implies a very large model set but we demonstrate how to obtain inference using some simple algebra and fast computation. Third, the proposed methodology is demonstrated with an empirical investigation of the long-run equilibrium relationships implied by the DSGE model. Important in this model are the long run responses of investment prices and productivity to technology shocks and that technology follows stochastic rather than deterministic trends.

The structure of the paper is as follows. In Section 2 the important features of the economic model used by Fisher (2006) are outlined. In Section 3 the basic econometric models of interest in this paper are introduced, including characterizations of the features implied by the economic model. We present priors, likelihood and the sampling scheme used in Section 4 together with the tools for inference in this paper including the posterior predictive probabilities (Geweke 1996 and Geweke and Amisano 2011) of alternative model features and the Laplace approximation. The posterior and predictive evidence of model features are presented in Section 5 as are estimates of important functions of parameters. In Section 6 we summarize conclusions and discuss possibilities for further research.

2 Long-Run Relations from a DSGE Business Cycle Model.

In this section we outline the features of a DSGE model that is based upon the real business cycle model of Fisher (2006), which is in turn based upon the competitive equilibrium growth model of Greenwood, Hercowitz, and Krusell (1997). We impose two simplifications: capital is not separated into equipment and structures; and technologies are given stochastic rather than deterministic trends. The general model was developed in Kydland and Prescott (1982) and detailed in King, Plosser and Rebelo (1988), and an interesting early econometric analysis is provided in King, Plosser, Stock and Watson (1991, hereafter referred to as KPSW). The reader is directed to these papers for the development of the model as we focus upon certain features that imply restrictions upon our reduced form econometric model that we want to weight using BMA.

The model suggests that a system of consumption, C_t, investment, X_t, and output, $\mathcal{Y}_t = C_t + X_t$, will share a balanced growth path since each is driven by shocks to two technologies: an investment specific technology, V_t; and neutral technology, A_t. We denote the logs of C_t, X_t, and $\mathcal{Y}_t$ by c_t, x_t, and $\mathbf{y}_t$ respectively.

The resource constraint and Cobb-Douglas production technology are given by

$$C_t + X_t \leq A_t K_t^{\lambda} H_t^{1-\lambda}, \quad 0 < \lambda < 1$$

and period $t+1$ capital stock is given by

$$K_{t+1} \leq (1-\delta) K_t + V_t X_t, \quad 0 < \delta < 1.$$

Fisher (2006) specifies technology as having stochastic rather than deterministic trends. The log of investment-specific technology, $v_t = \ln(V_t)$, and the log of neutral technology, $\mathrm{a}_t = \ln(A_t)$, are assumed to be independent simple random walks, possibly with drifts. In the empirical analysis we evaluate the evidence on the importance of deterministic and stochastic trends as well as the relative contribution to business cycle volatility of investment specific and neutral technology shocks.

The production technology and the resource constraint imply that we can represent the log real price of an investment good in consumptions goods by $p_t = -v_t$ and $p_t = p_{t-1} - \nu - z_{I,t}$. Since $\nu \geq 0$, this is in accordance with the downward trend we see in the price of an investment good. Neutral technology evolves by the process $\mathrm{a}_t = \gamma + \mathrm{a}_{t-1} + z_{N,t}$ where $\gamma \geq 0$ and $(z_{I,t}, z_{N,t})'$ has zero mean and constant covariance matrix.

A first implication of this model is that the variables c_t, x_t, and $\mathbf{y}_t$ will all be integrated of order one due to a common stochastic trend given by $\omega a_t + (1-\omega) p_t$ and the differences between any two will be stationary. This is not an unusual result in the balanced growth literature (see, for example, KPSW) and it implies that we can treat the Great Ratio relations $c_t - \mathbf{y}_t$ and $x_t - \mathbf{y}_t$ as valid cointegrating relations.

Denote by $h_t = \ln(H_t)$ the log number of hours worked which is assumed to have no unit root. Although it may have a trend over short periods, and the evidence suggests this, it is not possible for hours worked per capita to have a permanent trend. The log price of an investment good, p_t, and labour productivity, $a_t = \ln(\mathcal{Y}_t/H_t) = \mathbf{y}_t - h_t$, are assumed to have unit roots but p_t should not cointegrate with the other variables. Since h_t is assumed to be $I(0)$, and c_t, $\mathbf{y}_t$ and x_t are all assumed to be $I(1)$ sharing a common stochastic trend, the above assumptions imply that a_t will be $I(1)$ and $c_t - \mathbf{y}_t$ and $x_t - \mathbf{y}_t$ will be $I(0)$. The assumptions of Fisher (2006) preclude the above $I(0)$ relations having deterministic trends and this is not a feature we would expect to find over long samples. By allowing for structural breaks we may find trends in one or more subsamples, perhaps as a trend in the second period off-sets the effect of the trend in the first period, but we would expect that they are inherently temporary features.

Two important final restrictions apply. First, Fisher (2006) assumes that the long run response of p_t to $z_{I,t}$ will be nonzero, in fact negative, but its long run response to all other shocks will be zero. Second, the long run response of a_t to both $z_{I,t}$ and $z_{N,t}$ will be nonzero, but the long run response of a_t to any other shock will be zero. These restrictions identify the investment-specific technology shock, $z_{I,t}$, and the neutral technology shock, $z_{N,t}$.

3 A Set of Vector Autoregressive Models.

When a VAR process cointegrates, the model may be written in the vector error correction model (VECM) form. The VECM of the $1 \times n$ vector time series process $y_t = (p_t, a_t, h_t, c_t, x_t)$, $t = 1, \ldots, T$, conditioning on $l+1$ initial observations is

$$\Delta y_t = y_{t-1}\beta\alpha + d_t\mu + \Delta y_{t-1}\Gamma_1 + \ldots + \Delta y_{t-l}\Gamma_l + \varepsilon_t \tag{1}$$

where $\Delta y_t = y_t - y_{t-1}$. The $1 \times n$ vector of errors ε_t are assumed to be $iidN(0, \Omega)$. The matrices Γ_j $j = 1, \ldots, l$ are $n \times n$ and β and α' are $n \times r$ and assumed to have rank $r \in \{0, 1, ..., 5\}$. We define the deterministic terms $d_t\mu$ below.

As the model in the previous section implies two common stochastic trends, from the two technology shocks, there will be three cointegrating relations; that is $r = 3$. The model is quite specific about the form of these cointegrating relations and these impose overidentifying restrictions on the cointegrating vectors. Specifically, the model says that: p_t has a unit root but does not cointegrate with the other variables in the system; that hours worked, h_t, are stationary; and the great ratios of consumption to income and investment to income, $c_t - \mathbf{y}_t = c_t - a_t - h_t$ and $x_t - \mathbf{y}_t = x_t - a_t - h_t$, are stationary. We index the overidentifying restrictions by o and when all of the restrictions in the previous sentence are imposed we set $o = 2$ and this can only occur when $r = 3$. When no restrictions are imposed we set $o = 0$

which can occur for any r in $\{0, 1, ..., 5\}$. When we only impose the restriction that p_t has a unit root but does not cointegrate then $o = 1$. The restriction $o = 1$ is allowed to hold for $1 \leq r \leq 4$.

We allow for five lag lengths, $l \in \{2, 3, 4, 5, 6\}$. The deterministic processes are denoted by $d \in \{1, 2, 3, 4, 5\}$ and these processes are the five most commonly used combinations of trends in y_t and trends in $y_t\beta$ (for details see Johansen 1995, Section 5.7). The two empirically important deterministic processes in this paper are $d = 2$ which implies a linear trend in $y_t\beta$ and a linear drift in y_t, and $d = 4$ which implies no drift in the levels and no trend in the cointegrating relations. Some models implied by the deterministic processes will be observationally equivalent. For example, if $r = 0$ then different models of trends in $y_t\beta$ will be observationally equivalent. The treatment of *a priori* impossible and observationally equivalent models is explained in the next section when the prior is outlined.

Finally, the long run restriction to identify the technology shocks is employed. As discussed in the previous subsection, this restriction implies that the long run response of p_t is nonzero only for the investment-specific technology shocks and that the long run response of a_t is nonzero only for the investment-specific technology shock and the neutral technology shocks. This restriction can be parameterized using the standard Beveridge-Nelson form of the Wold representation of the VECM as

$$\Delta y_t' = C\Omega^{1/2} z_t' + C^* (L) \Omega^{1/2} \Delta z_t' \text{ where } C = \beta_\perp (\alpha_\perp \Gamma \beta_\perp)^{-1} \alpha_\perp .$$

The first two elements of $z_t = \Omega^{-1/2}\varepsilon_t$ are the investment-specific technology shock, $z_{I,t}$, and the neutral technology shocks, $z_{N,t}$, respectively. Further, $\Gamma = I_n - \Sigma_{i=1}^{l}\Gamma_i$. These restrictions imply the first two rows of $C\Omega^{1/2}$ will have the following zero entries[5]:

$$\begin{bmatrix} * & 0 & 0 & & 0 \\ * & * & 0 & \cdots & 0 \end{bmatrix}. \tag{2}$$

This restriction is obtained by the appropriate choice of $\Omega^{1/2}$ such that $\Omega = \Omega^{1/2}\Omega^{1/2\prime}$ and $C\Omega^{1/2}$ has the structure shown above. For an explanation of how such restrictions are implemented see the Appendix, and see Chang and Schorfheide (2003) and Del Negro and Schorfheide (2010) for further examples and discussion.

Since C has rank $n - r$, and $\Omega^{1/2}$ must have rank n, the zero restrictions on $C\Omega^{1/2}$ and the assumption of nonzero responses of p_t and a_t stated above imply C must have at least rank two, then this identification scheme can only apply if $r \in \{0, 1, \ldots, n-2\}$. This is consistent with the two technology shocks entering the system as stochastic trends.

Fisher (2006) assumes a break date around 1982 Q3, although our results suggest a slightly later date. We allow for a range of 16 break dates from the first quarter of 1982 until the last quarter of 1985. In our sample these periods cover the observations from $t = 137$ to $t = 152$. We index these dates by $\tau \in \{\underline{\tau}, \underline{\tau}+1, ..., \underline{\tau}+15, \overline{\tau}\}$ where the first break is at 1982Q1 ($\underline{\tau} = 137$), the second break occurs at 1982Q2, and so on

[5]The asterisks (*) imply no restriction is imposed.

until the last break at 1985Q4 ($\overline{\tau} = 152$).

In total we average using 255 models before the structural break, and 255 after the break and twelve break dates. This gives a total of over $780,000$ models that need to be computed. Computation of marginal likelihoods for this many multivariate models would be computationally challenging. As our interest is in the date when the break occurs and modelling after the break date, we need only estimate 510 marginal likelihoods for 12 break dates; a total of 6120 marginal likelihoods. This is still a large number of computations, but much more readily achievable. We provide further details on how this is achieved in Section 4.2.

4 Priors, Posteriors and Model Averaging.

In this section the priors and resultant posterior are presented. For notational convenience we collect the lag parameters into a $k_i \times n$ matrix $\Phi = [\Gamma_1' \quad \cdots \quad \Gamma_l']'$ and vectorize into $\phi = vec(\Phi)$. Conditional upon β, the model in (1) is linear in the equation parameters μ, $vec(\alpha)$ and ϕ, while conditional upon μ, $vec(\alpha)$ and ϕ, the model is linear in β. This fact makes it relatively straightforward to elicit priors on the parameters.

4.1 The Prior.

All models included for the averaging are treated equally likely. The set of models included for the averaging is a subset of the full set of models that result from all combinations of r, o, d and l and 16 break dates τ. We exclude impossible and meaningless models and, if two or more models are observationally equivalent, we include only one of them. To avoid notational burden, and because we use the same prior for models before and after the structural break, in this section we do not distinguish between parameters and models before and after the break dates except where necessary.

For Ω we use a proper inverted Wishart prior with scale matrix $\underline{S} = I_n 10$ and degrees of freedom $\underline{\nu} = n + 1$ as this prior is rather uninformative. The parameters in Φ are given a normal prior with zero mean and covariance matrix $1/\eta \underline{V}_\phi$ where $\underline{V}_\phi = (Iu + \underline{V}_0 (1-u))(I(1-u) + \underline{V}_1 u)$ and $u \in \{0, 1\}$ with prior probabilities $\Pr(u = 1) = \Pr(u = 0) = 0.5$. Here $\underline{V}_0 = \Omega \otimes I_{k_i}$ and $\underline{V}_1$ is the Litterman (1980, 1986) type prior for a VECM similar to that specified in Villani (2001). We use a gamma prior with mean $E(\eta) = 5$ and a relatively large variance $V(\eta) = 16.67$ for η. The settings we use provide a reasonable degree of shrinkage towards zero which has been shown to improve estimation (see Ni and Sun 2003).

We specify a weakly informative normal proper prior for $vec(\alpha, \mu)'$ conditional upon (Ω, β, M_i) (and hyperparameters discussed below) with zero mean and covariance matrix $1/\eta \underline{V}_a$ where $\underline{V}_a = \Omega \otimes I_r$.[6] Further details on the specification of the full prior is given at the end of the next subsection. The matrix β is given a Normal prior with zero mean and covariance matrix $n^{-1} I_{nr}$. The theory underpinning this specification is provided in Koop, León-González and Strachan (2010). Its motivation is to induce a uniform prior on the cointegrating space and a fast, efficient sampling scheme.

Let $a = \left(vec(\alpha)', vec(\mu)', \phi'\right)'$, $b = vec(\beta)$ and $\underline{V}$ be the block diagonal matrix with $\underline{V}_a$ and $\underline{V}_\phi$ on the diagonals . Introduce θ as the vector containing the elements of β, a, and Ω. The full prior distribution for the parameters in a given model is then

$$p(\theta, \eta, u|M_i) \propto \exp\left\{-\eta/2a'\underline{V}^{-1}a - n/2b'b\right\} n^{nr/2}$$
$$\times |\Omega|^{-(\underline{\nu}+n+1+r+uk_i)/2} \exp\left\{-1/2tr\Omega^{-1}\underline{S}\right\} \eta^{\frac{n(k_i+r)+1}{2}} \exp\left\{-5\eta/6\right\}.$$

Ideally all models would be treated as *a priori* equally likely, however this is not a straightforward issue in VECMs.[7] The priors for the individual elements of $i = (r, o, d, l)$ are not independent, as certain combinations are either impossible (such as when $r = n$ and $o = 2$), meaningless (such as, for example, $r = 0$ with $o = 1$) or observationally equivalent to another combination (such as the models with $r = n$

[6] If an informative prior is used on for the cointegrating space then we recommend the prior for α described in Koop, León-González and Strachan (2008). Further details on the development of this prior are available in an earlier working paper version of this paper: Strachan and van Dijk (2010).

[7] The authors are grateful to Geert Dhaene, John Geweke and an anonymous referee for useful comments on this issue.

and $d = 1$ or 2). The prior probability for impossible and meaningless models is set to zero. However, the researcher must carefully consider how she wishes to treat observationally equivalent models. Treating these models as just one model and then assigning equal prior probabilities to all models biases the prior weight in favour of models with $0 < r < n$. This could shift the posterior weight of evidence in favour of some economic theories for which we wish to determine the support.[8] Alternatively, these could be treated as separate models. A choice must be made and in this paper, observationally equivalent models are treated as one model.

A referee raised the interesting question as to whether it is appropriate to specify independent priors for r, o, d and l. One might expect, for example, that a strong deterministic process such as $d = 1$ might reduce the prior expectation of finding stochastic trends in the processes. This might imply that the probability $\Pr(r < n|d)$ may decrease as d increases. Similarly a shorter lag length, l, might be associated with a higher prior probability of finding (more) stochastic trends. We do not pursue this idea further, but note that it might be a worthwhile topic for investigation.

4.2 Posterior Analysis.

We conduct the empirical investigation in two stages. We first compute the estimated model probabilities and the timing of the structural break. We then estimate

[8] This issue could be viewed as a conflict between the desire to be uninformative across statistical models and the desire to be uninformative across economic models.

the functions of interest such as impulse responses and variance decompositions. We estimate the model probabilities using a Laplace approximation of predictive densities, and the impulse responses and variance decompositions are estimated using a Gibbs sampler.

When investigating the evidence on structural breaks we encounter two issues that need to be addressed: the strength of the prior information; and proliferation of models to compute. As we have a separate prior for the parameters before and after the break, we have effectively doubled the amount of prior information in the posterior and halved the amount of data (on average) used to estimate the parameters. This introduces the problem that the prior information is strong relative to the data. We could make the priors less informative, but then that increases their influence in the computation of the posterior probabilities (see discussion on Bartlett's paradox in Geweke 2005, Section 2.6.2). This situation implies a trade off between prior uncertainty about the parameters and posterior uncertainty about the models. To mitigate this issue we compute predictive probabilities (see, for example, Geweke 1996 and Geweke and Amisano 2011) from predictive densities.

The predictive densities can be derived for each model via the expression

$$p_i^{T_0,T_2} = p\left(y_{T_1+1}^{T_2}|y_{T_0}^{T_1}, M_i\right) = \frac{\int_\Theta p\left(y_{T_1+1}^{T_2}|y_{T_0}^{T_1}, M_i, \rho\right) p\left(y_{T_0}^{T_1}|M_i, \rho\right) p\left(\rho|M_i\right) d\rho}{\int_\Theta p\left(y_{T_0}^{T_1}|M_i, \rho\right) p\left(\rho|M_i\right) d\rho} \tag{3}$$

where $\rho = (\theta, u, \eta)$ and $y_{t_0}^{t_1}$ is the data from observation t_0 to t_1. For a model prior to the structural break at time τ, then $T_0 = 1$, $T_0 < T_1 < T_2$ and $T_2 = \tau - 1$, whereas for a model after the structural break, $T_0 = \tau$ and $T_2 = T$. As $T_2 \to \infty$ or $T_1 \to 0$, then $p\left(y_{T_1+1}^{T_2}|y_{T_0}^{T_1}, M_i\right)$ becomes the standard marginal likelihood. Thus the full predictive density for a process that switches from model M_i to model M_j at time τ is the product $p_i^{\tau} = p_i^{1,\tau-1} p_j^{\tau,T}$.

We estimate the numerator and denominator in (3), and therefore $p_i^{T_0,T_2}$, using the Laplace approximation. This technique approximates the integral by a second order expansion around the mode of the log of the integrand (see Tanner 1993). Although often described as a normal approximation, the Laplace approximation has been shown to work very well with many non-normal distributions. For example, Strachan and Inder (2004) apply the approach to the VECM where the integral is over linear subspaces which have bounded supports and non-standard distributions. Denoting by Ξ the set of models and using the notation $p^{T_0,T_2,\Xi} = \sum_{j \in \Xi} p_j^{T_0,T_2}$, then at a given break date τ we can compute the posterior predictive probability of a model M_i holding before (or after) the break using the $p_i^{T_0,T_2}$ as $p\left(M_i|y_{T_0}^{T_2}, \tau\right) = p_i^{T_0,T_2} \big/ p^{T_0,T_2,\Xi}$.

While using the Laplace approximation greatly speeds up the computation of the model probabilities, as discussed earlier, we allow the entire model structure to change at the break. This implies we have a very large number of models because the process is allowed to switch from any model before the break to any model after the break. To further reduce computation time, we make use of the fact that we are only interested in computing the probability that a break occurs at a point in time, and not which model after the break followed a specific model before the break. This observation greatly reduces the number of necessary computations.

If we have K possible break dates, then the posterior probability of a break at date $\tau = \tau^*$ is obtained by $p(\tau = \tau^*|y) = \left(p^{1,\tau^*-1,\Xi} p^{\tau^*,T,\Xi}\right) / \Sigma_{\tau=1}^{K} \left(p^{1,\tau-1,\Xi} p^{\tau,T,\Xi}\right)$. This expression, explained in a working paper (Strachan and van Dijk 2012), reduces the required number of computations from 780,300 to 6,120. The probability of a particular feature (e.g., $d = 2$) holding after the break is then obtained by summing the products of the probabilities of the models with that feature and the probabilities of their break dates.

To estimate the impulse responses we require draws from the posteriors of the models with non-negligible posterior weight. An expression for the posterior distribution of the parameters for any model given the data is obtained by combining the prior, $p(\theta, \eta, u|M_i)$, with the likelihood for the data $L(y|\theta, M_i)$ where y represents all data. That is,[9] $p(\theta, \eta, u|M_i, y) \propto p(\theta, \eta, u|M_i) L(y|\theta, M_i)$. As the sampler uses a Gibbs sampling scheme, it is necessary to present the conditional posterior for each parameter.

In the following results, we gather together terms to keep expressions notationally concise. Collect $y_{t-1}\beta$ and the vector $z_{2,t} = (d_{2,t}, \Delta y_{t-1}, \ldots, \Delta y_{t-l})$ into the vector $z_t = (y_{t-1}\beta, z_{2,t})$ and recall the definition of $a = \left(vec(\alpha)', vec(\mu)', \phi'\right)'$. As the model

[9] Note that as η and u are hyperparameters, they do not enter the likelihood.

is linear conditional upon $b = vec(\beta)$, standard results show that the posterior for a conditional on all other parameters will be normal with mean $\overline{a}$ and covariance matrix $\overline{V}$ constructed as

$$\overline{a} = \overline{V}\left(\Omega^{-1} \otimes I_{k_i+r}\right) vec\left(\Sigma_{t=1}^{T} z_t' \Delta y_t\right)$$

and

$$\overline{V} = \left[\left(\Omega^{-1} \otimes \Sigma_{t=1}^{T} z_t' z_t\right) + \eta \underline{V}^{-1}\right]^{-1}.$$

Next, the posterior for b conditional upon the other parameters will be normal with mean $\overline{b}$ and covariance matrix $\overline{V}_b$ which are constructed as

$$\overline{b} = \overline{V}_b\left(\alpha\Omega^{-1} \otimes I_n\right) vec\left(\Sigma_{t=1}^{T} y_{t-1}'\left(\Delta y_t - z_{2,t}\Phi\right)\right)$$

and

$$\overline{V}_b = \left[\left(\alpha\Omega^{-1}\alpha' \otimes \Sigma_{t=1}^{T} y_{t-1}' y_{t-1}\right) + n I_{nr}\right]^{-1}.$$

The posterior for η will be Gamma with degrees of freedom $\overline{\nu}_\eta = n\left(k_i + r\right) + 3$ and mean $\overline{\mu}_\eta = 1/\left(a'\underline{V}^{-1}a + 5/3\right)/\overline{\nu}_\eta$ (see, for example, Koop 2003). Finally, u will have a Bernoulli conditional posterior distribution with $\overline{p} = \Pr\left(u = 1|a, \Omega, \beta, y\right)$ equal to

$$\overline{p} = \exp\left\{-\eta/2a'\underline{V}_0^{-1}a\right\}/\left[\exp\left\{-\eta/2a'\underline{V}_0^{-1}a\right\} + \exp\left\{-\eta/2a'\underline{V}_1^{-1}a\right\}\right].$$

We use the following scheme at each step q to obtain draws of $(a, \Omega, \beta, \eta, u)$:

1. Initialize $(\Omega, b, a, \eta, u) = \left(\Omega^{(0)}, b^{(0)}, a^{(0)}, \eta^{(0)}, u^{(0)}\right)$;

2. Draw $\Omega|b, a, \eta, u$ from $IW\left(\underline{S} + u\eta A'A + \sum_{t=1}^{T} \varepsilon_t' \varepsilon_t, T + uk_i + r\right)$;

3. Draw $a|\Omega, b, \eta, u$ from $N\left(\overline{a}, \overline{V}\right)$;

4. Draw $b|\Omega, a, \eta, u$ from $N\left(\overline{b}, \overline{V}_b\right)$;

5. Draw $\eta|\Omega, b, a, u$ from Gamma$(\overline{\mu}_\eta, \overline{\nu}_\eta)$;

6. Draw $u|\Omega, b, a, \eta$ from $Bernoulli\left(\overline{p}\right)$;

7. Repeat steps 2 to 6 for a suitable number of replications.

To end this section we outline how we implement Bayesian model averaging. Let $\zeta\left(\rho, M_i\right)$ be an economic object of interest which is a function of the parameters, ρ, for a given model, M_i. Examples include estimates of impulse responses, forecasts, or loss functions. To report the unconditional (of parameters or any particular model) expectation of this object it is necessary to estimate $E\left(\zeta|y\right) = \sum\limits_{i \in \Xi} E\left(\zeta|y, M_i\right) p\left(M_i|y\right)$ where $E\left(\zeta|y, M_i\right)$ is the expectation of $\zeta\left(\rho, M_i\right)$ from model i marginal of ρ. Estimate $E\left(\zeta|y, M_i\right)$ as follows. First denote the q^{th} draw of the parameters from the posterior distribution for model M_i as $\rho^{(q)}$ and so the q^{th} draw of $\zeta\left(\rho, M_i\right)$ as $\zeta\left(\rho^{(q)}, M_i\right)$. Using J draws of the parameters from the posterior distribution for each of the models in Ξ, obtain estimates of $E\left(\zeta|y, M_i\right)$ from each model by $\widehat{E}\left(\zeta|y, M_i\right) = \frac{1}{J}\Sigma_{q=1}^{J}\zeta\left(\rho^{(q)}, M_i\right)$. These estimates are then averaged as $\widehat{E}\left(\zeta|y\right) = \sum\limits_{i \in \Xi} \widehat{E}\left(\zeta|y, M_i\right) \widehat{p}\left(M_i|y\right)$ in which $\widehat{p}\left(M_i|y\right)$ is an estimate of the posterior model probability, $p\left(M_i|y\right)$.

5 Evidence on Model Features.

In this section we provide empirical evidence on the support for the features of the DSGE model with two technology shocks and the various restrictions that this economic model implies for the econometric model. We begin with providing evidence on the timing of the structural break, and the posterior probabilities of the features of the reduced form VECM that are implied by this DSGE model. We estimate the posterior probability of a structural break occurring at a range of dates and, in subsequent analysis, focus upon the results for the data after the structural break. We then report the estimates of objects of interest including impulse response functions.

The variables and the data: The variables of interest are: log real price of an investment good measured in consumptions units, p_t; log labour productivity, a_t; log number of hours worked, h_t; log of consumption, c_t; and log investment, x_t. The data, which are seasonally adjusted, start in the first quarter of 1948 and end in the second quarter of 2009. Where appropriate, the data are measured in 1996 dollars deflated using a chain-weighted index of consumption prices. For more background on the data we refer to the working paper Strachan and Van Dijk (2012).

Break Dates and Stability of Great Ratios: Three break dates receive measurable support, 1984Q3, 1984Q4 and 1985Q1 with probabilities of 0.01%, 57.78% and 42.21% respectively. Of the 765 models over these three dates, there were 62 post break models with measurable support and 12 pre-break models with measurable support. From the pre-break models, the model with one cointegrating relation, no overidentifying restrictions, no drift and two lags, $(r, o, d, l) = (1, 0, 4, 2)$, with a break at 1984Q4 has 57.8% of the probability mass and the same model with with a break at 1985Q1 has 41.3% posterior probability. Thus, the models prior to the break provide strong evidence against the features we would expect under the economic model. While the results prior to 1984 have important information in them, there is evidence of several breaks in the 1970. As the aim of this paper is not a comprehensive treatment of all breaks, the remainder of the discussion focuses on the post break results as this period is well represented by a single, stable model.

Among the post break models, one model accounted for half the posterior mass and the top five models accounted for 99.73% of the posterior probability mass, as presented in Table 1. The posterior probabilities of the top five models are presented in Table 1. Post the break, the model $(3, 2, 2, 2)$ with a break at 1984 Q4 and the model $(2, 1, 4, 3)$ with a break at 1985 Q1 capture 79% of the probability mass. Although relatively few models get any support, it is clear that support is fairly strong for the top five models.

As the primary aim of this paper is to investigate the degree and role of model uncertainty in the empirical evidence on technology shocks and important features of DSGE models, we discuss the results with reference to a number of important papers in the literature. In Table 2 we have summarized the relevant features of the model by Fisher and compare this with results found in the literature in the KPSW study and the study by Centoni and Cubadda (2003, hereafter CC). We note that none of these specific models receive notable posterior support with our data set.

The posterior probability of having only two stochastic trends ($r = 3$) is 76%, although there is evidence of a third if the break is delayed until 1985 Q1. The DSGE model is driven by technology shocks which, in Fisher and KPSW, are stochastic trends. These are the only stochastic trends described in the economic model and KPSW assume a unique technology is (in the three variable model) the only stochastic trend that enters the system. The economic model of Fisher implies there are only two common stochastic trends. CC report evidence of an extra stochastic trend in a three variable system, but they then choose use the single trend model for inference. We conclude the evidence on an extra unit root is an empirical issue. It is possible the extra stochastic trend could be entering from the hours worked variable, h_t. While h_t does not display a trend, it does display a very large and slow cyclical component with peak-to-peak cycles of around 10 years duration and a sudden drop in 2009. Such large cycles and outliers are not easily captured in linear models (such as the VECM) and may be manifesting as evidence of a third unit root. An alternative explanation is provided in Chang, Doh and Schorfheide (2007), that hours may appear nonstationary if labour adjustment costs are not explicitly incorporated into the model.

The data are positively informative about the form of the cointegrating relations.

That is, there is a 76% probability that the Great Ratios are stable and that the investment price has a unit root and does not cointegrate with the other variables ($o = 2$). The assumptions that the price of an investment good is nonstationary and does not cointegrate with any other variable in the system, with 99% probability mass on $o = 1$ and $o = 2$, have very strong support. Since the joint posterior probability of ($r = 3, o = 2$) is 76%, and the posterior probability of ($r = 3, o = 0$ or $o = 1$) is zero, the evidence that the Great Ratios are stationary is positive, although not compelling.

According to the model of Fisher, it might be reasonable to expect few deterministic processes in the system and so to see $d \leq 3$ as the technologies are commonly described as random walks possibly with drifts, but the economic model does not suggest we would expect trends in the cointegrating relations ($d = 2$) or quadratic trends in the variables ($d = 1$). Table 3 presents the marginal probabilities of the various features of the VECM. There is a 76% probability that there is a drift in the levels and a trend in the cointegrating relations ($d = 2$) and 24% probability of no drift in levels and no trend in the equilibria ($d = 4$). With a 76% posterior probability that the cointegrating relations are the Great Ratios of consumption to income and investment to income, the trend may be picking up the decline in the savings that occurred since the mid 1990s.

Overall the evidence in the estimated probabilities for the features of the econometric model suggested by the DSGE model of Fisher (2006) is positive to strong.

But, with 25% of the mass on features not supported by the economic model, the evidence is not decisive and there remains considerable model uncertainty.

Table 1: Posterior probabilities, $P(M_i|y)$, of the top five models.

r	o	d	l	*Break Date*	$P(M_i\|y)$	Cumulative probabilities
3	2	2	2	1984.4	0.5552	0.5552
2	1	4	3	1985.1	0.2343	0.7895
3	2	2	2	1985.1	0.0968	0.8863
3	2	2	3	1985.1	0.0884	0.9747
3	2	2	3	1984.4	0.0226	0.9973

Table 2: Features of related models

	r	o	d	l
KPSW	2	2	3	8/6[10]
CC	2	0	3	1
Fisher[11]	0	0	3	3

[10] KPSW use 6 lags when testing for cointegration but 8 lags to produce the variance decompositions. In our results labelled KPSW below, we average over 6 to 8 lags.

[11] There is a distinction between the structure of the economic model of Fisher and the econometric model estimated by Fisher. The economic model of Fisher implies $r = 2$ and $o = 3$. However, the econometric model he estimates does not encompass all features of the economic model or use the same set of variables as our study. Further, we are assuming the model of Fisher used $d = 3$ as this

Table 3: Posterior marginal probabilities

of model features after the break.

$r=2$	$r=3$	$o=0$	$o=1$	$o=2$
0.2364	0.7638	0.0019	0.2348	0.7635
$d=2$	$d=4$	$d=5$	$l=2$	$l=3$
0.7636	0.2347	0.0019	0.6546	0.3454

Business Cycle Volatility due to Investment Specific and Neutral Technology Shocks: An important area of interest in DSGE models is the dynamics of $\mathbf{y}_t$, c_t, and x_t, including the role of the investment specific and neutral technology shocks in the business cycle. By decomposing the variance into the components due to these technology shocks, it is possible to gain an impression of the relative importance of these effects for the variability of the consumption, investment and output. As the model set includes models with the same features (r, o, d, and l) as those used in other studies, specifically KPSW and CC, it is possible to compare results across models used in other studies. KPSW and CC did not consider the two technology shocks *per se*, rather the role of permanent (possibly technology) shocks and transitory shocks. As the model in this paper has two additional variables (p_t and h_t), the results will differ from those if we had used exactly the models in KPSW and CC unless (p_t, h_t) is strongly exogenous to $(c_t, x_t, \mathbf{y}_t)$. KPSW and CC use output, $\mathbf{y}_t$,

was not clear from the paper.

whereas this paper uses productivity, $a_t = \mathbf{y}_t - h_t$. As a_t is a linear function of h_t which is also included in the model, the decomposition for $\mathbf{y}_t$ can be readily obtained from the estimation output.

KPSW derive an identification scheme for a decomposition based upon a single productivity shock entering these variables. This model is extended in Fisher to permit two types of permanent shocks, however in both cases the economic model implies that the Great Ratios ($c_t - \mathbf{y}_t$ and $x_t - \mathbf{y}_t$) will be stationary. As discussed above, results from our study suggest there is some uncertainty associated with this aspect of the theory as the evidence suggests there may be more than two stochastic trends entering the system. However, the equilibrium relations appear to be well described by the Great Ratios. Notwithstanding this ambiguity, it is not evident that the excess of stochastic trends affects estimates of other outputs such as proportions of the variance over the business cycle that can be attributed to the technology shocks.

Our interest is in the proportion of business cycle fluctuations due to the technology shocks in total, and the investment specific shocks, $z_{I,t}$, and neutral technology shocks, $z_{N,t}$, specifically. Therefore we adapt the approach of CC who consider the variance decomposition within the frequency domain.

Figure 1 presents the posterior distribution averaged over all models, of the proportion of the variance of c_t, h_t and $\mathbf{y}_t$ over the business cycle due to investment specific technology shocks and Figure 2 presents the posterior distribution of the proportion due to neutral technology shocks constructed by averaging over all models. These estimates are obtained from 30,000 draws from the posterior of each model. The plots show an amount of mass at zero for all three variables suggesting a slightly larger role for non-technology shocks. Relatively, the investment specific technology shocks are far more important than the neutral technology shocks and the proportions have considerable mass away from zero. The proportion of variation due to neutral technology shocks has little mass above 15%. We computed the same posterior densities as those reported in Figures 1 and 2 but using the CC, KPSW and the best models, as well as for all models in which the Great Ratios are the cointegrating relations. These estimates all share similar forms and pairwise comparisons showed no stochastic dominance of any distribution over any other.

Table 4 reports estimates of the mean proportions of the variance over the business cycle due to $z_{I,t}$ and $z_{N,t}$ for a range of models and model sets. The single models considered are those used by CC, KPSW and the best model in our model set. The model sets we consider are, first, the set of all models and, second, the set of all models in which the Great Ratios are stable. Comparing across different models and model sets, the proportions appear relatively robust to specification. The CC model is the only one that does not include the Great Ratios as stable cointegrating relations and we see that this results in slightly higher proportions of the variance due to investment specific technology shocks for output and consumption, although the difference is not large. Imposing the Great Ratios (as in KPSW, the best model and the set of all models with stable Great Ratios) slightly reduced the role of $z_{I,t}$ in business cycle variation for consumption and output, but not for investment. Overall, the technology shocks explain between 26% and 38% of the variation over the business cycle for these variables.

Relatively speaking, the investment specific technology shocks seem far more important than the neutral technology shocks as they explain between 70% and 97% of the total variation due to technology shocks.

Table 4 reports the 95% credible intervals for the variance decompositions for the models and model sets we consider.[12] An interesting point that arises from a study of these intervals is that allowing for model uncertainty does not necessarily increase the degree of uncertainty about the variance decompositions. The last two columns of this table report results for sets of models: all models; and all models with balanced growth restrictions. In many cases the credible intervals in these columns are narrower than those for the single model estimates in the first three columns. We attribute this effect, through correlations among outputs, to something analogous to the well know variance reduction from diversification in an investment portfolio of financial assets.

[12] We are grateful to Frank Schorfheide for suggesting the inclusion of these intervals and to investigate the implications of model uncertainty and parameter uncertainty in the posterior distributions of the variance decompositions.

Propagation of Shocks: We conclude by discussing the responses in investment price, hours, output, consumption and investment to both investment specific and neutral technology shocks. The percentiles of the posterior distribution of these responses are shown in the panels in Figure 3. The left panels show responses to an investment specific shock, $z_{I,t}$, while the right panels show responses to a neutral technology shock, $z_{N,t}$. Included in each panel is the mean impulse responses to the technology shocks constructed from 30,000 draws of the parameters in each included model. It is immediately apparent that most of the densities have mass near zero. Note that these plots incorporate much more uncertainty than the usual ones that condition upon a model and so cannot be interpreted in the same way. The densities in Figure 3 reflect both parameter and model uncertainty and so any intervals will naturally be much wider and so can lead to different conclusions. Unlike with the variance decompositions, we do not benefit from model averaging in estimation of impulse responses.

The mean path of investment price p_t to shocks in $z_{I,t}$ and $z_{N,t}$ are very much as would be expected from the model of Fisher and agree with the form of his simulated and estimated responses. The plot of the posterior response of p_t to an investment specific shock is persistently negative, while the response to a neutral technology shock shows only a slightly positive response, again agreeing with the results of Fisher. Output has a positive response to both shocks and the sizes of the responses are similar in magnitude to those of Fisher, with a stronger and more persistent response to $z_{N,t}$. Neither consumption nor investment display the persistent or large responses suggested by Fisher's simulations. In contrast to the simulations of Fisher, investment has an initially negative, then positive then declining response to investment specific technology shocks. Investment simply declines in response to neutral technology shocks. The response of hours worked differs depending upon the type of technology shock in that investment specific shocks lead to an eventual fall in hours worked while the neutral shocks have a positive and persistent effect. The magnitudes of the responses in hours is slightly smaller than those of Fisher. Both our results and those of Fisher suggest persistence in the response, however, our responses to investment specific shocks differ from those reported in Fisher (2006) in an important respect. Our results show a fall in hours worked while Fisher's results suggest an increase in hours worked.

While many of the mean paths agree with those of Fisher (2006), the extra uncertainty leads to an important difference in the conclusions. With the exception of the response of investment price to $z_{I,t}$, the mean paths are not far from zero if we take the full posterior as the metric for distance. The posterior credible intervals for many of the responses encompass the origin but these intervals do not capture all of the information on the densities. The densities for the responses of the investment price to a neutral technology shock in Figure 4 at three horizons give an impression of the general form and evolution of the response densities for other variables. The leptokurtic form and mass generally around zero is representative of the responses of a_t, $\mathbf{y}_t$, h_t, c_t and x_t to investment specific shocks, and of h_t to neutral technology shocks.

The responses of a_t, $\mathbf{y}_t$, c_t and x_t to neutral technology shocks show a very different response. As an example, Figure 5 presents the posterior distribution of the response of output $\mathbf{y}_t$ to an neutral technology shock at three different horizons. These densities all have platykurtic forms for responses after one or two periods. However, they become increasingly bimodal at the time since the shock increases. This bimodality is not due to different models producing different paths as the responses for the best model in Table 1 produced very similar responses. A number of possible explanations exist, and one is that the identifying restrictions, at least for $z_{N,t}$, disagree with the data. However, as we are using just identifying restrictions it is not possible for us to report evidence at this point on this question.

Table 4: Estimated proportion of variance over the business cycle due to the investment-specific technology shock, $z_{I,t}$, and the neutral technology shocks, $z_{N,t}$. 95% Credible intervals are shown in parentheses.

	CC model	KPSW model	Best model	All models	Balanced Growth models
			Consumption - c_t		
$z_{I,t}$	0.299	0.179	0.212	0.223	0.212
	(0.02,0.75)	(0.01,0.53)	(0.02,0.55)	(0.02,0.58)	(0.02,0.55)
$z_{N,t}$	0.050	0.018	0.041	0.038	0.040
	(0.00,0.20)	(0.00,0.04)	(0.00,0.15)	(0.00,0.13)	(0.00,0.14)
			Investment - x_t		
$z_{I,t}$	0.254	0.320	0.233	0.227	0.234
	(0.02,0.62)	(0.01,0.79)	(0.02,0.60)	(0.02,0.59)	(0.02,0.60)
$z_{N,t}$	0.051	0.01	0.035	0.033	0.035
	(0.00,0.20)	(0.00,0.01)	(0.00,0.13)	(0.00,0.12)	(0.00,0.13)
			Output - $\mathbf{y}_t$		
$z_{I,t}$	0.281	0.160	0.202	0.216	0.202
	(0.02,0.69)	(0.01,0.44)	(0.02,0.53)	(0.02,0.56)	(0.02,0.53)
$z_{N,t}$	0.097	0.045	0.082	0.076	0.081
	(0.02,0.29)	(0.02,0.08)	(0.02,0.22)	(0.02,0.20)	(0.02,0.22)

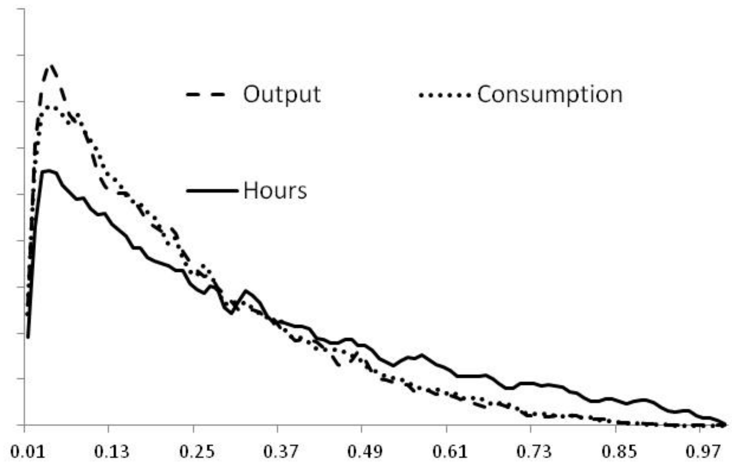

Figure 1: Posterior densities of the proportion of the variance over the business cycle of output, consumption and hours that is due to investment specific technology shocks, $z_{I,t}$.

6 Conclusions.

This paper presents a Bayesian model averaging approach in order to investigate the empirical support for several features of the DSGE type of real business cycle model of Fisher (2006) which we subject to two types of technology shocks. An important component of this model is the restrictions of long run responses that are used to identify investment specific and neutral technology shocks. For many of the features implied by this model we find reasonable support and some, such as stability of the Great Ratios, receive quite strong support, although not before 1984. Further, the

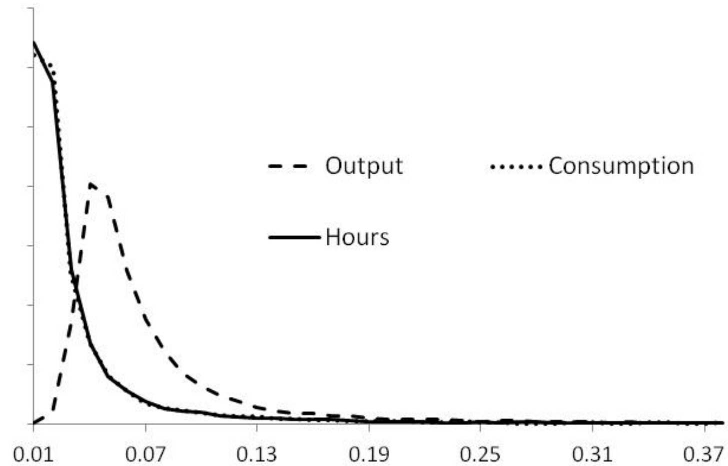

Figure 2: Posterior densities of the proportion of the variance over the business cycle of output, consumption and hours that is due to neutral technology shocks, $z_{N,t}$.

impulse responses demonstrate that the predictions of the model are quite plausible although there is considerable uncertainty around these estimates.

The methodology is an important contribution of this paper. The approach results in unconditional inference on these features of the vector autoregressive model as the effect of any one model on the inference has been averaged out, and so model uncertainty is incorporated into the analysis. Techniques are developed for estimating marginal likelihoods for models defined by such features as the number of stable relations (cointegration rank), overidentifying equilibrium restrictions, deterministic processes and short-run dynamics. To account for structural breaks and model switching, methods are presented to make inference computationally feasible in a very large set of multidimensional models.

The methods presented in this paper, without the structural breaks, have already found applications in several other areas. Koop, Potter and Strachan (2005) investigate the support for the hypothesis that variability in US wealth is largely due to transitory shocks. They demonstrate the sensitivity of this conclusion to model uncertainty. Koop, León-González and Strachan (2008) develop methods of Bayesian inference in a flexible form of cointegrating VECM panel data model. These methods are applied to a monetary model of the exchange rate commonly employed in international finance. Other current work includes investigating the impact of oil prices on the probability of encountering the liquidity trap in the UK and stability of the money demand relation for Australia.

We end with mentioning two topics for further research. First, although our mixing over priors partially addresses this issue, there remains the issue of the robustness of the results with respect to wider prior and model specifications. Very natural extensions of the approach in this paper are to consider forms of time variation in the model itself as Cogley and Sargent (2001, 2005) and Primiceri (2005) do for the VAR. These papers suggest there is considerable variation in the parameters, particularly over the 1970s. However, the reduced rank restrictions due to cointegration introduce further conceptual and computational issues for time varying models and a first example of how to implement such a time varying VECM is presented in Koop, León-González and Strachan (2011). Alternatively, in using a SVAR for business cycle analysis one may use prior information on the length and amplitude of the period of oscillation (see Harvey, Trimbur and van Dijk 2007). Unresolved issues in this work include systematic use of inequality conditions which imply a more intense, or better, use of MCMC algorithms. Extending this approach to large model sets is yet to be considered and, given the computational issues, remains a challenge. Second, one may use the results of our approach in explicit decision problems in international and financial markets like hedging currency risk or evaluation of option prices.

Author Affiliations: Rodney W. Strachan (Corresponding author), Research School of Economics, The Australian National University. Strachan is also a Fellow of The Centre for Applied Macroeconomic Analysis. Herman K. van Dijk Econometric Institute, Erasmus University Rotterdam and Econometrics Department VU University Amsterdam. Van Dijk is an Honorary Fellow of the Tinbergen Institute. Both authors are Senior Fellows of the Rimini Centre for Economic Analysis.

7 Appendix

Identification of the technology shocks: In this appendix we detail how the two technology shocks are identified by the zero restrictions on $C\Omega^{1/2}$ shown in Section 3. Let $\widetilde{\Omega}^{1/2}$ denote any unique decomposition of Ω such that $\Omega = \widetilde{\Omega}^{1/2}\widetilde{\Omega}^{1/2\prime}$. This could be the unique Cholesky decomposition into a lower triangular matrix or taken from the square root of the singular value decomposition. As we use the singular value decomposition below we will give this definition for $\widetilde{\Omega}^{1/2}$. Take the singular value decomposition of the $n \times n$ matrix A as $A = USV'$ where $U \in O(n)$, $V \in O(n)$ where $O(n)$ is the orthogonal group of $n \times n$ matrices, and $S = diag(s_1, s_2, \ldots, s_n)$ where $s_i \geq 0$. A square root of A may be obtained by the construction $A^{1/2} = US^{1/2}V'$.

To identify z_t we need to choose an orthogonal matrix $\widetilde{U}$ such that $\Omega^{1/2} = \widetilde{\Omega}^{1/2}\widetilde{U}$ such that $\Omega = \Omega^{1/2}\Omega^{1/2\prime} = \widetilde{\Omega}^{1/2}\widetilde{U}\widetilde{U}'\widetilde{\Omega}^{1/2\prime} = \widetilde{\Omega}^{1/2}\widetilde{\Omega}^{1/2\prime}$.

Let c^1 be $(1 \times n)$ vector of the first row of $C\widetilde{\Omega}^{1/2}$ and let c^2 be the $(1 \times n - 1)$ of the second to last elements in the second row of $C\widetilde{\Omega}^{1/2}$. Next let the projection matrix projecting into the space of c^1 be $P^1 = c^{1\prime}\left(c^1 c^{1\prime}\right)^{-1} c^1$. Similarly, let the projection matrix projecting into the space of c^2 be P^2.

Then $\widetilde{U} = U^1 U^2$ where U^1 is the orthogonal matrix spanning the column space of the singular value decomposition of $P^1 = U^1 SV'$. Further,

$$U^2 = \begin{bmatrix} 1 & 0 \\ 0 & U^{22} \end{bmatrix}$$

where U^{22} is the orthogonal matrix spanning the column space of the singular value decomposition of $P^2 = U^{22}SV'$. We can then define z_t as $z_t = \Omega^{-1/2}\varepsilon_t$.

8 References.

Bates, J. M., and C. W. J. Granger, "The Combination of Forecasts," *Operational Research Quarterly* **20** (1969), 451–468.

Centoni, M. and G. Cubbada, "Measuring the business cycle effects of permanent and transitory shocks in cointegrated time series," *Economics Letters* 80 (2003), 45-51.

Chang, Y., and F. Schorfheide, "Labor-Supply Shifts and Economic Fluctuations," *Journal of Monetary Economics* **50(8)** (2003), 1751-1768.

Chang, Y., Doh, T., and F. Schorfheide, "Non-stationary Hours in a DSGE Model," *Journal of Money, Credit, and Banking*, 39(6) (2007), 1357-1373.

Cogley, T., and T. J. Sargent, "Evolving Post-World War II US Inflation Dynamics," *NBER Macroeconomics Annual* **16** (2001), 331–373.

Cogley, T., and T. J. Sargent, "Drifts and Volatilities: Monetary Policies and Outcomes in the Post WWII U.S.," *Review of Economic Dynamics* **8** (2005), 262-302.

Del Negro, M., and F. Schorfheide, "Bayesian Macroeconometrics" in: J. Geweke, G. Koop, and H.K. van Dijk, ed., *Handbook of Bayesian Econometrics* (2011), Chapter 7, pp. 293-389.

Diebold, F.X. and J. Lopez, "Forecast Evaluation and Combination," in G.S. Maddala and C.R. Rao (eds.), *Handbook of Statistics* (Amsterdam: North-Holland 1996), 241-268.

Draper, D., "Assessment and propagation of model uncertainty (with discussion)," *Journal of the Royal Statistical Society Series B* **56** (1995), 45-98.

Fernández, C., E. Ley, and M. Steel, "Model uncertainty in cross-country growth regressions," *Journal of Applied Econometrics* **16** (2001), 563-576.

Fisher, J. D. M. "The Dynamic Effects of Neutral and Investment-Specific Technology Shocks," Manuscript, Fed. Reserve Bank Chicago 2005.

Fisher, Jonas D. M. "The Dynamic Effects of Neutral and Investment-Specific Technology Shocks," *Journal of Political Economy* **114** (2006), 413-451.

Geweke, J. "Bayesian reduced rank regression in econometrics," *Journal of Econometrics* **75** (1996), 121-146.

Geweke J. *Contemporary Bayesian Econometrics and Statistics* (Wiley: Hoboken, NJ, 2005).

Geweke, J., and G. Amisano, "Hierarchical Markov Normal Mixture Models with Applications to Financial Asset Returns," *Journal of Applied Econometrics* **26** (2011), 1-29.

Greenwood, J., Z. Hercowitz, and P. Krusell, "Long-Run Implications of Investment-Specific Technological Change," *The American Economic Review* **87** (1997), 342-62.

Harvey, A.C., T.M. Trimbur and H. K. van Dijk, "Trends and cycles in economic time series: A Bayesian approach," *Journal of Econometrics* **140** (2007), 618-649.

Hendry, D. F. & M. P. Clements, “Pooling of Forecasts,” *Econometrics Journal* **5** (2002), 1-26.

Hodges, J., “Uncertainty, policy analysis and statistics,” *Statistical Science* **2** (1987), 259-291.

Hoogerheide, L.F., R. Kleijn, F. Ravazzolo, H.K. Van Dijk, and M. Verbeek, “Forecast Accuracy and Economic Gains from Bayesian Model Averaging using Time Varying Weights,” *Journal of Forecasting* **29** (2010), 251-269.

Johansen, S., *Likelihood-based Inference in Cointegrated Vector Autoregressive Models* (New York: Oxford University Press, 1995).

King, R., C.I. Plosser, and S. Rebelo, “Production, growth, and business cycles I: The basic neoclassical model,” *Journal of Monetary Economics* **21** (1988), 195-232.

King, R.G., C.I. Plosser, J.H. Stock, and M.W. Watson, “Stochastic trends and economic fluctuations.” *The American Economic Review* **81** (1991), 819-840.

Koop, G., *Bayesian Econometrics* (Wiley, Chichester, 2003).

Koop, G., R. León-González and R. W. Strachan, “Bayesian inference in a cointegrating panel data model” *Advances in Econometrics, Volume* **23** (2008).

Koop, G., R. León-González and R. W. Strachan, “Efficient posterior simulation for cointegrated models with priors on the cointegration space” *Econometric Reviews* **29** (2010), Issue 2, 224-242.

Koop G., R. León-González and R. W. Strachan, "Bayesian Inference in the Time Varying Cointegration Model," *The Journal of Econometrics* **165** (2011), 210-220.

Koop G., R. León-González and R.W. Strachan, "A Reversible Jump algorithm for inference using instrumental variables" forthcoming in *The Journal of Econometrics* (2012).

Koop, G. and S. Potter, "Bayesian Analysis of Endogenous Delay Threshold Models," *Journal of Business and Economic Statistics* **21** (2003), 93-103.

Koop, G., S. Potter and R.W. Strachan, "Re-examining the Consumption-Wealth Relationship: The Role of Model Uncertainty," *Journal of Money, Credit and Banking* **40** (2005), No. 2–3, 341-367.

Kydland, F. and E. Prescott, "Time to build and aggregate fluctuations," *Econometrica* **50** (1982), 173-208.

Leamer, E., *Specification Searches* (New York: Wiley, 1978).

Litterman, R. B., "A Bayesian procedure for forecasting with vector autoregressions," Massachusetts Institute of Technology, Mimeo (1980).

Litterman, R. B., "Forecasting with Bayesian vector autoregressions — Five years of experience," *Journal of Business and Economic Statistics* **4** (1986), 25–38.

McConnell, M. and G. Perez-Quiros, "Output Fluctuations in the United-States: What Has Changed Since the Early 1980's," *American Economic Review* **90** (2000), 1464–1476.

Min, C. and A. Zellner, "Bayesian and non-Bayesian methods for combining models and forecasts with applications to forecasting international growth rates," *Journal of Econometrics* **56** (1993), 89-118.

Newbold, P. & D. Harvey, "Tests for Multiple Forecast Encompassing," *Journal of Applied Econometrics* **15** (2001), 471–482.

Ni, S. X. and D. Sun, "Noninformative Priors and Frequentist Risks of Bayesian Estimators of Vector-Autoregressive Models," *Journal of Econometrics* **115** (2003), 159-197.

Poirier, D., *Intermediate Statistics and Econometrics: A Comparative Approach* (Cambridge: The MIT Press, 1995).

Primiceri, G., "Time Varying Structural Vector Autoregressions and Monetary Policy," *Review of Economic Studies* **72** (2005), 821-852.

Raftery, A. E., D. Madigan, and J. Hoeting "Bayesian model averaging for linear regression models," *Journal of the American Statistical Association* **92** (1997), 179–191.

Sala-i-Martin, X., G. Doppelhoffer, and R. Miller, "Determinants of long-term growth: A Bayesian averaging of classical estimates (BACE) approach," *American Economic Review* **94** (2004), 813-835.

Stock, J. H. and M. W. Watson, "Has the Business Cycle Changed and Why?," *NBER Macroeconomics Annual*, Volume 17 (MIT Press, 2002).

Strachan, R. W. and B. Inder, "Bayesian Analysis of The Error Correction Model," *Journal of Econometrics* **123** (2004), 307-325.

Strachan, R. W. and H. K. van Dijk, "Evidence on a DSGE Business Cycle Model subject to Neutral and Investment-Specific Technology Shocks using Bayesian Model Averaging," *CAMA Working Paper* 3/2012 (2012).

Tanner, M. A., *Tools for statistical inference : methods for the exploration of posterior distributions and likelihood functions.* (2nd Edition). (Springer Series in Statistics, New York: Springer, 1993).

Terui, N. and H. K. van Dijk, "Combined forecasts from linear and nonlinear time series models," *International Journal of Forecasting* **18**(3) (2002), 421-438.

Villani, M., "Bayesian prediction with cointegrated vector autoregressions," *International Journal of Forecasting* **17** (2001), 585-605.

Wright, Jonathon H., "Bayesian Model Averaging and exchange rate forecasts," *Journal of Econometrics* **146** (2008), 329-341.

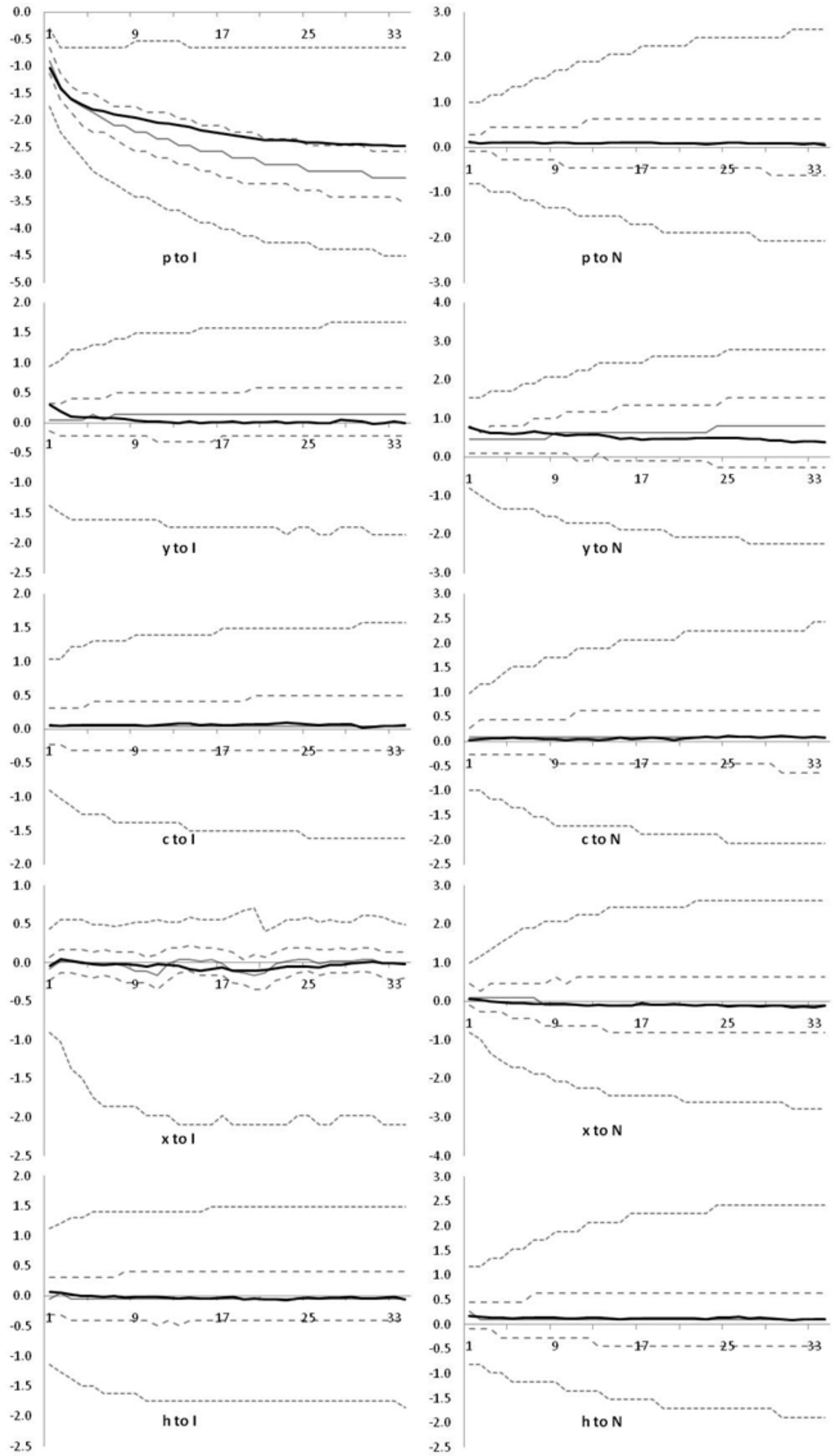

Figure 3: The 20%, 40%, 50%, 60% and 80% percentiles of the posterior distribution

of the response of investment price, p, output, y, consumption, c, investment, x, and hours, h, to investment specific, I, and neutral, N, technology shocks. The x-axis is periods since the shock and the y-axis is in percent.

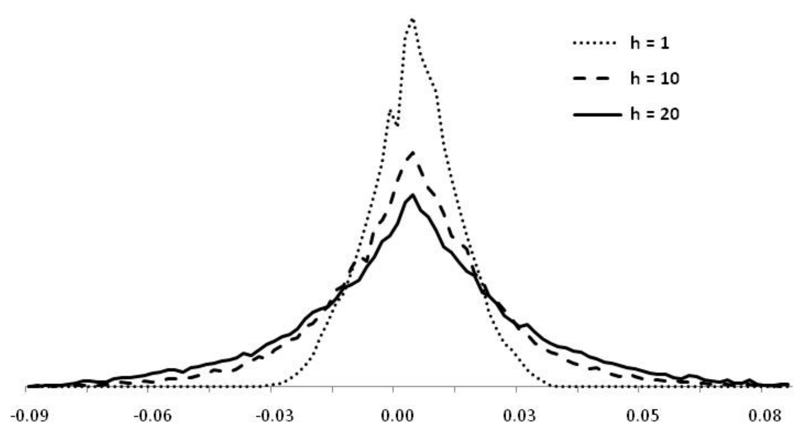

Figure 4: Posterior densities of the impulse responses of investment price, p_t, to a neutral technology shock, $z_{N,t}$, at horizons h = 1, h = 10 and h = 20. The x-axis is in percent.

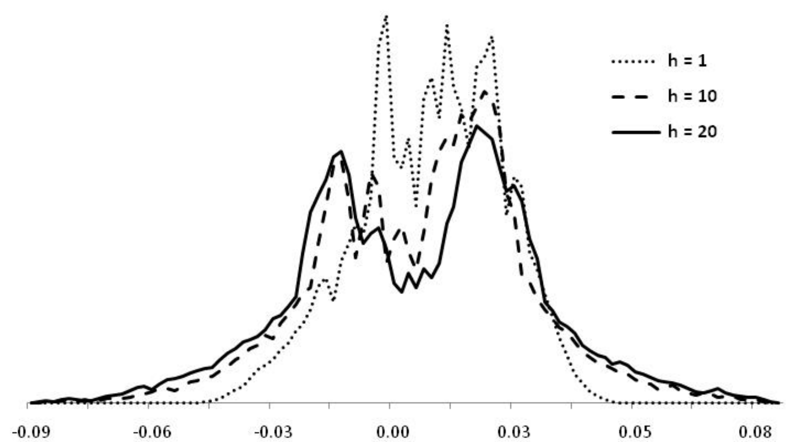

Figure 5: Posterior densities of the impulse responses of output to a neutral technology shock, $z_{N,t}$, at horizons h = 1, h = 10 and h = 20. The x-axis is in percent.