A Bayesian VAR-GARCH Analysis of Sectoral Labour Reallocation

November 20, 2000

Gianluigi Pelloni*
Department of Economics
Faculty of Political Science
University of Pisa
Via Serafini, 3
Pisa, 56126
Italy
(gigi@specon.unipi.it)

Wolfgang Polasek**
Department of Statistics
and Econometrics
University of Basel
Holbeinstrasse, 12
CH-4051 Basel
Switzerland
(Wolfgang.Polasek@unibas.ch)

Abstract

The paper proposes to model sectoral and aggregate employment dynamics by a multivariate GARCH-in-mean (VAR-GARCH-M) model. We investigate the question of the direction of volatile shocks between sectoral employment shares and total employment growth. We show that the VAR-GARCH-M model is an appropriate extension of the VAR-based approach to the analysis of the aggregate effects of reallocation shocks. Using a Bayesian approach we show that the VAR-GARCH-M structure explains employment growth behaviour much better than the simple VAR model. We find that sectoral shocks in the U.S. labour market account for more than 60% of the variance of total employment growth within a VAR-GARCH-M framework. Moreover, sectoral shocks in the US labour market account for more than 60% of the variance of total employment growth within a VAR-GARCH-M framework.

JEL Classification Numbers: C11, E24, E32

Keywords: Intersectoral Labour Reallocation; Sectoral Shifts; MCMC methods; Non-linearities; Bayes Factors; VAR-GARCH Models; Impulse Response Functions.

1 Introduction

In recent years, a major point in the discussion of dynamic sectoral employment models has centered around the Lilien or sectoral shift hypothesis. In brief, this hypothesis claims that the inter-sectoral variance of employment growth rates is a measure of labour reallocations and is an important source for the flow of shocks from the sectors to total employment.
Lilien's paper (1982a) started an increased interest for studies on the macro-economic effects of reallocation shocks in labour market models. Lilien (1982a, p. 780) suggested to use an index variable - now also called Lilien index - in the following way: 'σ_t is a measure of the dispersion of employment demand conditions throughout the labor market. Shocks to the economy that have differing impacts on firms, such as a rise in oil prices, will lead to an increase in σ_t' . For the empirical study Lilien proxied the index σ_t through a cross-sectional variance. Subsequently, the index was critisized because sectoral shocks could not be separated from aggregate shocks. Many studies have tried to overcome this "observational equivalence" problem which accompanies the Lilien hypothesis. Campbell and Kuttner (1996, CK henceforth) started a new way to model the relationship between aggregate and sectoral employment by multivariate time series models.
In this paper we will extend Lilien's concept by defining sectoral dispersion indices: we define $\sigma_{t,j}$ as a measure of the dispersion of employment demand conditions of the labor market in sector j. Shocks to the economy will lead to an increase in the sector-specific index $\sigma_{t,j}$. The index $\sigma_{t,j}$ is proxied in a sectoral employment model through an ARCH effect, i.e. a heteroskedastic sectoral variance component. This approach allows us to analyse the sectoral shift hypothesis within a multivariate dynamic volatility model in order to evaluate the macro-economic relevance of sectoral shocks.
Now the sectoral time series model is based on the idea that Lilien's dispersion index is reframed as sectoral volatilities in the class of VAR-GARCH-M models which is frequently used in financial econometrics. Univariate ARCH models have been introduced by Engle (1982, 1983) and multivariate ARCH models by Engle and Kroner (1995). Multivariate models analyse potential nonlinear sectoral shocks within a VAR structure. The price for such a modeling strategy is that dynamic interactions between the aggregate and the sectors have to be modeled by large dimensional parameter matrices. To overcome the small sample problems in large scale dynamic models we will

use Bayesian simulation techniques to carry out the necessary computations. We will extend the sectoral growth model of CK in two directions: First, we will show that sectoral shift models might be parameterized in terms of a multivariate GARCH-in-mean (or VAR-GARCH-M) model. In such a model we make the assumption that reallocation shocks might be proxied by the volatilities in a VAR model. Second, because of the huge amount of parameters we will use a Bayesian estimation method using the Markov Chain Monte Carlo (MCMC) methods for the vector GARCH model [(1)] and we will use Bayes factors for hypothesis testing. Third, we use the data set as in Mills, Pelloni and Zervoyianni, (1995, MPZ henceforth) for our employment model to see if our results can corroborate their findings.
The relevance of intersectoral labour reallocation as a source of aggregate employment fluctuations is at the centre of an ongoing research (for a survey c.f. Gallipoli and Pelloni, 2000). This debate persists because the problem of "observational equivalence" is deeply rooted in sectoral shifts analysis (Lilien 1982b; Abraham and Katz, 1986). There are two major economic arguments:
- A fall in aggregate demand may induce firms to lay off workers temporarily with consequent changes in aggregate employment and unemployment. On the other hand, sector-specific shocks, affecting the allocation of demand across sectors, could bring about intersectoral movements of workers which, because of the time-consuming processes of searching, retraining and relocating, could also alter the levels of (un)employment (Lilien, 1982a).
- If cyclical responsiveness varies across sectors, dispersion measures, instead of reflecting reallocation shocks, could be mirroring aggregate shocks. Therefore any positive (negative) correlation between unemployment (employment) and dispersion indices would reflect aggregate disturbances and not labour market turbulences. It follows that observationally equivalent predictions could be generated by at least two different models of business cycle shocks.
This paper's contribution to the existing literature is two-fold [(2)]. The impact of sectoral shifts on aggregate employment is explicitly introduced for the first time within a VAR framework by modeling shocks as ARCH processes. In doing so, we reformulate Lilien's hypothesis in this heteroskedastic framework for evaluating its macro-economic effects. Though we do not provide a theoretical economic model to support our empirical investigation, we think there is a compelling intuition behind it, i.e. the size of sectoral shocks and its persistence can affect economic behaviour. We hope that this

paper may operate as a basis for future research since new extensions of the Bayesian approach has been used: MCMC estimation, model selection and innovation analysis, and impulse response functions.
The paper is organized as follows. In the next section 1.1 we describe the economic framework that is used to model sectoral shocks non-linearities. We present in section 2 the formulation of employment growth model with ARCH effects and we finally introduce the VAR-GARCH-M model. In section 3 we discuss the Bayesian methodology employed both for estimation and hypothesis testing. We explain how the forecasting, the impulse response function and the variance decomposition procedures are implemented. In section 4 we present the data set together with the estimation results for the US labor market between 1975 and 1990. It contains a discussion of our new results and in particular the explanatory power of allocative shocks for aggregate employment variations. Section 5 concludes with final remarks.

2 Aggregate effects of sectoral shocks

Lilien's dispersion hypothesis (1982a) claims that, in an economy with limited mobility of resources across markets (labour in particular), changes in the composition of employment demand will trigger a process of job reallocation which will affect aggregate employment. Lilien's (1982a) original turnover framework explicitly appeals to labour search as the underlying economic mechanism and Lucas and Prescott (1974) is invoked as the theoretical reference model. Lilien makes no attempt to frame his hypothesis within a fully developed theoretical model. He proceeds directly to work out the essential skeleton of a turnover structure where intersectoral shifts in demand composition operate as the basic driving force of unemployment fluctuations. Under reasonably mild assumptions, he derives a reduced form unemployment equation characterised by a positive relationship between the unemployment rate and a measure of dispersion of employment demand conditions. For the empirical implementation Lilien proxies intersectoral dispersion with a weighted standard deviation of cross-sectoral employment growth rates(3). He finds a strong positive correlation between U.S. unemployment rate and his dispersion measure, and interprets this result as evidence in favour of the sectoral shifts hypothesis.
Lilien's

approach is open to severe criticism because of the inherent problem of "observational equivalence" pointed out in the introduction. Different approaches have been proposed to overcome this difficulty in the measurement of sectoral shifts. These lines of research fall essentially into two classes: that using micro/panel data (Murphy and Topel, 1987; Loungani and Rogerson, 1989; Starr-McCluer, 1993) and that using aggregate data. This latter approach can itself be divided into four further subclasses: that exploring the correlation between observed dispersion and the vacancy rate (Abraham and Katz, 1984; Davis 1987); the line of research based on constructing a dispersion index defined in terms of sectoral stock prices instead of sectoral employment (Loungani, Rush and Tave, 1990; Brainard and Cutler 1993); the "purging" methodology (Lilien, 1982b; Abraham and Katz, 1984; Samson, 1990; Neelin, 1987, Neuman and Topel 1991; Mills, Pelloni and Zervoyianni 1995, 1996), which is aimed at decomposing Lilien's dispersion proxy into a component measuring the sectoral response to aggregate shocks and one measuring sector- specific factors; finally the approach based on VAR systems free of dispersion measures and aimed at modelling sectoral shifts directly [(3)].

Prominent within this fourth subclass is the work by Campbell and Kuttner (1996, henceforth CK). CK constructs a VAR system including the growth rates of aggregate employment and of manufacturing employment shares, which they subsequently expand to include five more sectors. The underlying structure of their model is developed by imposing short-run and long-run identifying restrictions. The analysis of driving forces is structured around short-run and long-run identifying restrictions. The emerging outcomes vary quite widely depending on the underlying restrictions. Sectoral shocks account for 6% of total employment variance under the short-run triangular bivariate system; they can explain 82% of the variance under the long-run restriction within the seven dimensional system.

CK provides a linear model characterised by a "symmetric" response of aggregate employment to sectoral shocks: a positive shock to the manufacturing sector will increase aggregate employment growth and vice-versa for a negative shock. Thus, CK assumes that sectoral shocks operate like aggregate shocks, i.e. it is the shock's direction that counts in determining the aggregate response. This treatment of allocative disturbances somehow distorts the nature of sectoral shifts analysis [(5)]. In fact, as stressed by Davis (1986), sectoral shocks cannot be seen as having a "positive-negative" structure, but instead, they should be understood as either "favourable" or "unfavourable" to the existing allocation of resources. In order to analyse the aggregate impact of labour market turbulence it is more pertinent to look at the magnitude than at the direction of allocative shocks.

The stylised fact that the size of shocks matters independently of direction seems to suggest that sectoral shifts analysis should accommodate a non-linear framework. Since the essence of Lilien's hypothesis, as we will discuss in more detail in section 3, is embedded in his assumption of heteroskedasticity, we introduce non-linearity by modelling the volatility of sectoral shocks directly. We expand CK's approach to a multivariate non-linear time series model where non-linearities enter the model through the mean (ARCH-in-mean-feedback) and through the variance structure.

The possible significance of aggregate shocks variability has long been recognised. For instance the so-called Lucas proposition (Lucas, 1973) claims that the effects of unanticipated nominal disturbances vary inversely with these shocks' volatility. Engle (1982, 1983), in developing ARCH models, also explored as well the inverse relationship between inflation variability and real output. Thus the inclusion of aggregate shocks volatility into this sort of analytical framework should, at least in principle, be uncontroversial. Allowing for volatilities of sector-specific shocks seems to be an reasonable cogent assumption to make when analyzing the employment effects of macro-economic turbulences. It is the essence of the sectoral shifts hypothesis that unemployement/employement cyclical fluctuations are generated by variation in the variance of sectoral shocks. Thus modeling the variance of sectoral shocks is a necessary feature of sectoral shifts analysis. Its omission might lead to misspecifation of the time series models and a misleading summary of the dynamic response in such models.

Therefore the presence and correlation of volatilities will be relevant for the analysis of cyclical fluctuations and reallocations within the sectoral shifts models. Also, the sectoral shifts hypothesis claims that an uneven and idiosyncratic arrival process of information about the desired employment allocation across sectors would explain a large fraction of employment and unemployment variation. Because of this arrival pattern of news it is reasonable to expect that during a period of turbulence, shocks could exhibit a large volatility component.

Volatility effects imply that any information on changes in sector-specific fundamentals reaches a sector in clusters, so that these allocative shocks are responsible for a changing volatility structure as well. Our maintained hypothesis can be couched as follows. The sectoral shifts hypothesis claims that sector-specific shocks lead to changes in sectoral employment growth rates because labour and capital are, at least in part, sector specific. In our view the dynamics of labour reallocation can be affected not only by large reallocation shocks but also by their persistence. This means that large shocks of either sign will be followed by a large shock of either sign, and the fact that the direction of labour reallocation may or may not change frequently is not of primary concern. What really matters is the presumption that the macro-economic effects would emerge as a response to sizeable reallocation shocks or from persistent volatility.

3 Sectoral shocks and volatility models

In the following section we reformulate Lilien's sectoral shifts hypothesis in terms of sectoral time series models containing volatility effects: Lilien's model (1982a) is developed around a firm net hiring function characterized by a firm-specific component modeled as a random process with time-varying variance:

$$y_{t,j} = \mu_t + \varepsilon_{t,j}, \qquad t = 1, \ldots, T; \quad j = 1, \ldots, M \tag{1}$$

where $\varepsilon_{t,j} \sim N\left(0, \sigma_t^2\right)$. Ignoring employment quits and letting the behaviour of a specific sector be reflected by the behaviour of its typical firm, equation (1) can be interpreted as the employment rate of change at sectoral level, which can be decomposed into an aggregate component common to all sectors, μ_t, and a component, $\varepsilon_{t,j}$ specific to sector j. We can call (1) the sectoral growth model with heteroskedasticity. Given the state of econometrics at the time Lilien wrote the paper, it is surprising that he did not try to incorporate the assumption of heteroskedasticity explicitly in his model but used instead an intersectoral measure of labour mobility.

Todays methods allow to add a proper dynamic dimension to Lilien's hypothesis of heteroskedasticity by explicitly modeling the variance driving process as an ARCH component. The hypothesis contained in equation (1) can then be enriched by introducing the following assumptions about sectoral heteroskedasticity:

$$\varepsilon_{j,t} \;=\; u_{j,t}\sqrt{h_{t,j}}, \quad j = 1, \ldots, M, \quad t = 1, \ldots, T, \tag{2}$$

$$h_{t,j} = \quad \alpha + \textstyle\sum_{i=1}^{q} \theta_i^j \varepsilon_{j,t-i}^2, \quad \text{for } \alpha > 0, \quad \text{and } \theta_i^j \geq 0. \tag{3}$$

In this new specification of the model the sector-specific component is generated by equation (2), where $u_{j,t} \sim NiiD(0,1)$, $t = 1, \ldots, T$, and $h_{t,j} \equiv Var\left(\varepsilon_{j,t}|\Omega_{t-1}\right)$ is given by the process specified in (3), which is defined for the information set Ω_{t-1} available at time $t-1$. Thus the model suggests that the conditional variance, $h_{t,j}$, is a positive function of the past squared sectoral innovations, regardless of their signs, so that large errors tend to be followed by a large error and small errors by a small error. Modeling sectoral shocks as in (2) adds a new dimension to the model because it makes it possible to capture not only a time-varying variance but also a potential phenomenon of volatility clustering.
A further extension can be introduced by taking into account the possible feedback of the conditional variance on the conditional mean. This can be accomplished by respecifying equation (1) as

$$y_t^j = y_t + \gamma_0^j + \sum_{i=1}^{r} \gamma_i h_{t-i}, \qquad j = 1, \ldots, M. \tag{4}$$

This assumption would entail the insight that the behaviour of economic agents, in the presence of reallocation shocks, is affected not only by the size of these shocks but also by their volatile structure. For instance, risk averse agents could be inclined to move from high volatility sectors to low volatility ones. Equations (2), (3) and (4) provide then an ARCH-in-mean model for the employment growth rate of sector j.

3.1 The VAR-GARCH-M sectoral growth model

After having introduced the basic GARCH structure for volatile employment growth models we can now formulate the general multivariate VAR-GARCH-M model assuming normal distributions:

$$\mathbf{y}_t|\Omega_{t-1} \sim N(\mu_t, \mathbf{H}_t), \qquad t = 1, \ldots, T, \tag{5}$$

where $\mathbf{y}_t$ is a vector whose elements are the growth rates of aggregate employment and of sectoral employment shares. Thus we define a M-dimensional

VAR(k)-GARCH(p,q)-M(r) model as

$$\mathbf{y}_t = \mu_t + \varepsilon_t = \beta_0 + \sum_{i=1}^{k} \mathbf{B}_i \mathbf{y}_{t-i} + \sum_{i=1}^{r} \mathbf{\Psi}_i vech \mathbf{H}_{t-i} + \varepsilon_t, \tag{6}$$

$$vech \mathbf{H}_t = \alpha_0 + \sum_{i=1}^{p} \mathbf{A}_i vech \mathbf{H}_{t-i} + \sum_{i=1}^{q} \mathbf{\Theta}_i vech \left(\varepsilon_{t-i} \varepsilon^{'}_{t-i} \right). \tag{7}$$

Equation (6) is called the 'mean equation' and equation (7) is the 'variance equation'. Furthermore, $\mathbf{y}_t$ is a $(M \times 1)$ vector of time series; $\mathbf{H}_t$ is a $(M \times M)$ diagonal conditional variance-covariance matrix; $vech\mathbf{H}_t$ is a $(\tilde{M} \times 1)$ vector with $\tilde{M} = M(M+1)/2$ elements of variances and covariances; ε_t is M-dimensional process of mutually and serially uncorrelated random errors and so $vech(\varepsilon_t \varepsilon^{'}_t)$ is a $\tilde{M}$ dimensional vector; α_0 and β_0 are respectively $(\tilde{M} \times 1)$ and $(M \times 1)$ vectors of time invariant intercept coefficients. The $\mathbf{B}_i$'s are the VAR coefficient and the $\mathbf{\Psi}_i$'s are $(M \times \tilde{M})$ ARCH-in-mean coefficient matrices while the $\mathbf{A}_i$'s and $\mathbf{\Theta}_i$'s are $(\tilde{M} \times \tilde{M})$ GARCH coefficient matrices. $vech\mathbf{H}_t$ denotes the column stacking operator for the elements of a symmetric matrix on and below the main diagonal. For certain cases it might be sufficient to use in (6) instead of $vech\mathbf{H}_t$ only $\mathbf{h}_t$, where $\mathbf{h}_t$ is a M-dimensional vector of conditional variances (the main diagonal of $\mathbf{H}_t$);

According to our VAR-GARCH-M model the conditional means are functions of the contemporaneous and lagged values of the conditional variances so as to verify whether the information content of the conditional variances is relevant in determining the estimates of the conditional mean values. In turn each conditional variance depends upon the past values of the squared shocks, its own lagged values and the lagged values of the conditional variances relative to the other equations.

Since $y^j_t = \triangle \log(N^j_t/N_t)$ is the growth rate of employment in sector j we interpret the main diagonal elements of $\mathbf{H}_t$ in (5), the conditional variance of the sector specific shock, as a proxy of labour reallocation for sector j. We think that, consistently with Lilien's sectoral shifts hypothesis, these measures of sector-specific employment reallocation should enter the growth equation (6) and forms the 'ARCH-in-mean' part of the model. The extra dimension we are adding in here is the feedback of the conditional variances (a proxy of reallocations) to the VAR equation.

The volatilitvdvdtoral

experienced in the past, the larger the current volatility and its impact on behaviour. E.g. a large change in variance in this period will increase next period variance, thereby increasing the chance of a large shock in the next period. Though a general macro-economic framework is still lacking for these effects, it is reasonable to explore sectoral growth dynamics empirically. As stated in section 2, we think that the idiosyncratic informative structure underlying the sectoral shifts hypothesis and the workings of market mechanisms may justify the presence of volatility clustering.

Furthermore we wish to explore if volatility spillovers exist or if the conditional variance changes are sector-specific. Changes in conditional volatility could be brought about by shocks to sector-specific fundamentals without affecting other sectors volatility and mean values. Alternatively volatility changes in one sector could affect volatility changes in other sectors and influence their employment growth rates. Thus we will to model sectoral shocks characterized by a changing, clustering volatility whose effects may spillover to other sectors.

While Lilien's proxy was introduced to measure intersectoral employment dispersion and its macro-economic effects, our attention is focused on adding an extra dynamic dimension when modeling reallocative shocks directly and on verifying if and how much this new class of models matters. No experiment has been implemented yet in order to quantify the volatility of sector-specific shocks and its relevance in explaining changes in sectoral employment shares and aggregate employment fluctuations. Thus, a multivariate GARCH model for sectoral employment growths extends the sectoral shifts analysis and goes beyond it.

3.2 The prior assumptions

We intend to do extensive testing in the VAR-GARCH model and therefore it is necessary to specify a proper prior distribution. Since the number of parameters increases tremendously with the dimension of the system we tried to find a prior distribution which permits a quick specification but is flexible enough to carry out a sensitivity analysis with respect to the prior. Thus, we adopt the tightness approach by Litterman (1986) for the AR and the GARCH coefficients (in modified form) since this approach allows through the tightness parameters a convenient sensitivity analysis with respect to the prior distributions.

For the prior distribution of the regression parameters in (6) we have assumed a zero prior mean and a diagonal covariance matrix with tightness variances, i.e. the variances of lagged coefficients are inversely proportional to their lag orders and all covariances are set to zero. Let β_i^m be the m^{th} column $(m = 1, \ldots, M)$ of the $M \times M$ matrix $\mathbf{B}_i$, then the tightness prior assumes $\beta_i^m \sim N[\mathbf{0}, \mathbf{I}_M/i], i = 1, \ldots, k$. Similarly, if Ψ_i^m is the m^{th} column of $\mathbf{\Psi}_i$. Then we assume $\Psi_i^m \sim N[\mathbf{0}, \mathbf{I}_M/i], i = 1, \ldots, r$.
The prior distribution of the GARCH coefficients is constructed in a similar way as for the AR coefficients. We assume a tightness covariance structure as above but with a positive prior mean (0.05 or 0.01, depending on the number of parameters). Additionally, we restrict the GARCH coefficients to the stationarity region of the multivariate GARCH process. Engle and Kroner (1995) show that the multivariate GARCH(p,q) process is stationary if the eigenvalues of the sum of the GARCH coefficient matrices $\sum_{i=1}^{p} \mathbf{A}_i + \sum_{i=1}^{q} \mathbf{\Theta}_i$ in (7) is smaller than 1. For the GARCH matrices $\mathbf{A}_i$ and $\mathbf{\Theta}_i$ we assume a multivariate normal distribution with diagonal tightness covariance matrices:

$$\mathbf{A}_i \sim N[\mathbf{A}_*; c\mathbf{I}_{\tilde{M}^2}/i], \quad i = 1, \ldots, p; \tag{8}$$

$$\mathbf{\Theta}_i \sim N[\mathbf{\Theta}_*; c\mathbf{I}_{\tilde{M}^2}/i], \quad i = 1, \ldots, q; \tag{9}$$

where c is an appropriate chosen scaling factor (e.g. $c = 0.1$), which can be used for a sensitivity analysis. For the constant $\alpha_{\mathbf{0}}$ in (7) we assume a (positive) truncated normal distribution with mean 0.1 and variance 1. Other distributions (like a chi-square or an exponential) can also be used, but as long as the prior distribution is rather flat, the functional form is not important. The prior means for the GARCH coefficients are assumed to be small but positive depending on the size of the model: $\mathbf{A}_* = \mathbf{\Theta}_* = 0.01\mathbf{I}_{\tilde{M}}$. The proposal distribution in the Metropolis step is assumed to be a normal distribution where the parameters are obtained by 2-3 trial runs starting from a draw of the prior distribution; this is called 'iterative proposal'. This distribution is used as a candidate generator for the accept/reject algorithm in the MCMC chain and after an iterative start-up we found always satisfactory accept/reject ratios.
The pragmatic choice of the prior distribution helps to improve the convergence of the MCMC run and we found that for reasonable sample sizes the posterior distribution is quite close to the (classical) constrained likelihood function. The modes of the posterior density have been found close to the

maximum likelihood estimates for special cases where we could make the comparisons (see Liu and Polasek 1999). Also, we have made a sensitivity analysis with respect to the tightness prior by varying the scaling factor c for the prior covariance matrix. Again we find that the model selection with marginal likelihoods gives identical answers if c is varied over a reasonable range.

4 Test and Estimation Procedures

In this section we describe the Bayesian approach to test and estimate multivariate volatile time series models. A comparative recent presentation of Bayesian econometrics can be found in Poirier (1995) and details for the estimation procedure can be found in the BASEL package of Polasek (2000). First we show how Bayes test can be carried out by Bayes factors obtained from a MCMC output. Section 3.2 shows how unit root test can be calculated from 'fractional' marginal likelihoods. Finally we show how predictive distributions and impulse response function are obtained from a MCMC estimation.

4.1 Bayes factors and a rule of thumb

In order to select the appropriate order of the VAR process and to choose between the linear and non-linear versions of our model we will use Bayes factors [(6)]. When comparing any two models the Bayes factor (BF) can be calculated using the marginal likelihood concept. Let $\mathbf{y}$ denote the observed data and θ_j the parameters under model M_j, then the marginal likelihood is defined as

$$f(\mathbf{y}|M_j) = \int f(\mathbf{y}|\theta_j, M_j)\, f(\theta_j|M_j)\, d\theta_j. \tag{10}$$

If we denote by $f(\mathbf{y}|\mathrm{M}_1)$, the marginal likelihood for model 1, and by $f(\mathbf{y}|\mathrm{M}_2)$, the marginal likelihood for model 2, then the BF for M_2 versus M_1 is the ratio

$$\mathrm{BF}_{21} = \frac{f(\mathbf{y}|\mathrm{M}_2)}{f(\mathbf{y}|\mathrm{M}_1)}. \tag{11}$$

Using the usual rules for computing odds, we find the posterior probabilities for models M_1 and M_2 by assuming prior probabilities for the two models.

Note that if the two models are equally likely a priori, i.e. $\mathrm{P}(\mathrm{M}_1) = \mathrm{P}(\mathrm{M}_2) = 0.5$ the the BF is equal to the posterior odds ratio. In particular, one may want to use the so-called 9:19:99 rule for evaluating Bayes factors: $\mathrm{BF} > 9$ are remarkable, $\mathrm{BF} > 19$ are significant, and $\mathrm{BF} > 99$ are highly significant hypotheses. In practice it is convenient to use log-Bayes factors

$$\ln \mathrm{BF}_{21} = \ln f(\mathbf{y}|\mathrm{M}_2) - \ln f(\mathbf{y}|\mathrm{M}_1). \tag{12}$$

If M_1 and M_2 are any two models under consideration we can use the differences of the log marginal likelihoods to judge the importance of models. Now $\ln f(\mathbf{y}|\mathrm{M}_2)$ is the log marginal likelihood of model 2 and $\ln f(\mathbf{y}|\mathrm{M}_1)$ is the log marginal likelihood of model 1. The above cut-off points for the log BF are now: $\ln 9 = 2.2$, $\ln 19 = 2.9$, and $\ln 99 = 4.6$. Roughly speaking this means, we can use the numbers 2, 3 and 4 1/2 when comparing models with log marginal likelihoods. In practice we will compare any two GARCH-M time series models of different orders by Bayes factors where the marginal likelihoods are computed by the method of Chib and Jeliazkov (1999).
It should be noted that the BF in (9) and the log BF in (10) are only meaningful if the prior distribution is proper. However if the prior is improper then alternative methods to compute BF have been proposed [(7)]. In the next section we show how fractional Bayes factors (O'Hagan, 1995), which uses 'fractions' of the data to create a prior distribution can be used to test for unit roots without specifying a prior distribution explicitly [(8)].

4.2 Unit roots test with marginal likelihoods

Though not all Bayesian econometricians agree about pre-testing for unit roots [(9)], we proceed to test the univariate properties of the relevant time series using marginal likelihoods. In this section we outline how the classical augmented Dickey-Fuller (DF) regression for unit roots can be used to calculate the marginal likelihoods for a Bayes test.
Let y_t be the univariate time series we want to check for unit roots and let z_t be the first differences so that $z_t = y_t - y_{t-1}$. Following Dickey and Fuller's approach we consider the following autoregression models to test unit roots.
Δ-AR: The AR(p) model for first differences:

$$z_t = \alpha_1 z_{t-1} + \ldots + \alpha_p z_{t-p} + u_t, \tag{13}$$

DF-AR1: The plain Dickey-Fuller regression model:

$$z_t = \alpha_0 y_{t-1} + \alpha_1 z_{t-1} + \ldots + \alpha_p z_{t-p} + u_t. \tag{14}$$

DF-AR2: The Dickey-Fuller regression model with mean:

$$z_t = \mu + \alpha_0 y_{t-1} + \alpha_1 z_{t-1} + \ldots + \alpha_p z_{t-p} + u_t \qquad (14a),$$

DF-AR3: The Dickey-Fuller regression model with trend:

$$z_t = \mu + \beta t + \alpha_0 y_{t-1} + \alpha_1 z_{t-1} + \ldots + \alpha_p z_{t-p} + u_t \qquad (14b),$$

$$t = 1, \ldots, T; \qquad p = 1, \ldots, p_{max;}$$

Equation (13) defines the $\triangle -$ AR model model: if the time series is I(1) then the first differences should be a stationary AR(p) process. Equations (14), (14a) and (14b) define three alternative stationary models: the so-called Dickey-Fuller regression model (DF-AR1 model), Dickey-Fuller regression model with mean (DF-AR2 model) and Dickey-Fuller regression model with mean and trend (DF-AR3 model) respectively. We select the specification with highest probability according to the marginal likelihood criterion discussed in section 3.1. If the estimated marginal likelihood for the $\triangle$-AR model is higher than the marginal likelihoods of the alternative stationary models then the unit root hypothesis is supported. Of course if any of the alternative models is selected by the marginal likelihood, then the stationarity hypothesis is instead supported. Thus, using a fractional or an informative prior distribution, the marginal likelihoods for the DF regression models (13) to (14b) can be calculated and the best model together with the lag length will be chosen by looking at the highest marginal likelihood.
Using this methodology we tested the growth rate of aggregate employment and the growth rates of the employment shares of the durable, non-durable, transportation and services sectors. We can conclude from the tests (c.f. table 1) that the transformation to 4^{th} differences produces stationary time series ([10]). Previously, always applying the same methodology, we had tested the levels $\left(n_t^j = \log N_t^j\right)$ of these variables and found all of them to be $I(1)$. A further extension of this approach to unit root testing is given in Polasek and Ren (1999) where its was shown that outliers and break points can be included in the above testing procedures. Our overall conclusion from all these tests is that the employment growth rates are stationary.

4.3 Estimation of the VAR-GARCH-M model

The estimation of VAR-GARCH-M model was done by the BASEL package which is described in Polasek (2000) and the full conditional distribution for this model class can be found in Polasek and Ren (2000). Therefore we will give in this section only a brief outline of the estimation procedure.
Let $\tilde{\mathbf{B}}$ be all parameters of the mean equation in (6), then full conditional posterior distribution $p(\tilde{\mathbf{B}}|\mathbf{Y}, \tilde{\mathbf{A}})$ can be shown to be a multivariate normal distribution. Let $\tilde{\mathbf{A}}$ be all the parameters of the variance equation in (7), then the full conditional distribution $p(\tilde{\mathbf{A}}|\mathbf{Y}, \tilde{\mathbf{B}})$ is not given in closed form and we need a Metropolis step. The candidate distribution is a multivariate normal distribution where the mean is the old draw $\tilde{\mathbf{A}}^{old}$ from the previous iteration and the covariance matrix $\hat{\mathbf{D}}_h^m$ is obtained iteratively from smaller Metropolis runs starting from the prior distribution. After a few runs we calculate mean and variance matrix again and repeat the iterative proposal until a satisfactory accept/reject ratio is obtained.
The lag orders in the VAR(k)-GARCH(p,q)-M(r) model are determined by the maximum of the marginal likelihoods which are computed according to formula (10). Since we estimate the models by MCMC we can also use now the convenient methods of Chib (1995) and Jeliazkov and Chib (1999) to compute the exact marginal likelihoods. The implementation into the MCMC algorithm is described in Polasek and Ren (2000) and in the 'BASEL-package' (Polasek 2000).
A simulation study for a VAR-GARCH model was conducted in Liu and Polasek (1999). For the parameter estimates it was found that the VAR-GARCH model which is implemented in the BASEL-package gives the best results for the VAR-GARCH model. It turned out that the classical estimates which are obtained from commercial program packages could not even recover the right sign of the coefficients of the simulated processes. Furthermore, the standard deviations of the coefficients are surprisingly accurate even for a small number of iterations of the MCMC algorithm. The convergence and the MCMC is fast for small dimensions of the time series system and efficient, as can be seen easily from the autocorrelation function of the output trajectories. The implementation of the MCMC algorithm follows in two steps the block-at-a-time Metropolis - Hastings algorithm. Thus, the usual requirements of an irreducible and aperiodic stationary Markov chain are given and there are no constraints to move around in the feasible param-

eter space of a general VAR-GARCH-M model. Details on the convergence diagnostics and the MCMC sampling procedure can be found in the BASEL-package (see Polasek 2000) which is available on the web.

4.4 Forecasting the VAR-GARCH-M model

In order to generalize the concept of the impulse responses function (IRF) to non-linear time series systems we have first to show how to forecast the VAR-GARCH-M model. Based on the equation system of (6) and (7) we derive the predictive density for μ_t and $\mathbf{H}_t$ and in analogy to the linear IRF we show how the non-linear IRF for mean and variances are obtained numerically from the predicted first and second conditional moments.

4.4.1 Forecasting the conditional covariance matrices $\mathbf{H}_t$

Using conditional expectations E_t at time t for the error terms ε_t and the conditional matrices $\mathbf{H}_t$ we can calculate the future conditional covariance matrices (at time t for s step ahead and each sample point $l = 1, \ldots, L$)

$$\begin{aligned} vech\mathbf{H}_{t+s}^{(l)} &= \alpha_0^{(l)} + \sum_{i=1}^{q} \mathbf{\Theta}_i^{(l)} vech E_t(\varepsilon_{t+s-i}^{(l)} \varepsilon_{t+s-i}^{\prime(l)}) \\ &\quad + \sum_{j=1}^{p} \mathbf{A}_j^{(l)} vech\mathbf{H}_{t+s-j}^{(l)} \end{aligned} \tag{15}$$

with $vech\mathbf{H}_t$ denoting the half-vectorisation of $\mathbf{H}_t$ and the conditional mean is

$$E_t \varepsilon_{t+s}^{(l)} = \begin{cases} \hat{\varepsilon}_{t+s}^{(l)} = y_{t+s} - \mu_{t+s}^{(l)} & \text{for } s \leq 0, \\ 0 & \text{for } s > 0. \end{cases} \tag{16}$$

4.4.2 Forecasting future means

Having found the forecasts of the conditional covariance matrices we can now turn to predict the conditional means. The estimates of the s-step ahead forecasts from the MCMC simulation of size L are calculated as the mean of the predictive distribution, i.e.

$$\hat{\mathbf{y}}_{t+s} = \frac{1}{L} \sum_{l=1}^{L} \left[\beta_0^{(l)} + \sum_{i=1}^{k} \mathbf{B}_i^{(l)} E_t \mathbf{y}_{t+s-i} + \sum_{i=1}^{r} \mathbf{\Psi}_i^{(l)} vech\mathbf{H}_{t+s-i}^{(l)} \right] \tag{17}$$

with

$$E_t \mathbf{y}_{t+s} = \begin{cases} \mathbf{y}_{t+s} & \text{for } s \leq 0 \\ \hat{\mathbf{y}}_{t+s} & \text{for } s > 0. \end{cases} \tag{18}$$

4.5 The impulse response function (IRF)

Impulse response function are used in VAR systems to describe the dynamic behaviour of the whole system with respect to unit shocks in the residuals of the time series. For nonlinear time series systems like multivariate GARCH models the concept has to be extended to generalized impulse response function. In extension of the approach of Hamilton (1994, p. 318) and Koop et al. (1996) we define the generalized impulse response function to be the derivative

$$\partial \mathbf{y}_{t+s} / \partial \varepsilon_t^{'} = \mathbf{M}_s, \quad s = 1, 2, \ldots \tag{19}$$

for the VAR-GARCH-M model and each column of $\mathbf{M}_s$ is defined as numerical derivative in direction ε_{t+1}

$$\Delta \hat{\mathbf{y}}_{t+s}(\varepsilon_{t+1}) = \frac{1}{s}[E_t(\mathbf{y}_{t+s}|\varepsilon_{t+1}, \Omega_t) - E_t(\mathbf{y}_{t+s}|\Omega_t)], \quad s = 1, 2, \ldots \tag{20}$$

where Ω_t is the information set up to time t and ε_{t-1} varies over all unity vectors. The expectation is taken as the mean of the predictive distribution and is estimated by the average over the simulated future paths calculated from the MCMC output.

The difference between the predicted value of the vector $\hat{\mathbf{y}}_{t+s}$ at time $t+s$ in (20) corresponds to the j^{th} column of the matrix $\mathbf{M}_s$. By doing a separate simulation for impulses to each component of the innovation vector ($j = 1, \ldots, M$), all of the columns of $\mathbf{M}_s$ can be calculated, i.e.

$$\mathbf{M}_s = [\Delta \hat{\mathbf{y}}_{t+s}(\mathbf{e}_1), \ldots, \Delta \hat{\mathbf{y}}_{t+s}(\mathbf{e}_M)],$$

where $\mathbf{e}_1, \ldots, \mathbf{e}_M$ are the M unity vectors of order M. Note that the impulse response function of a nonlinear system is not time invariant, it depends on the time t, the forecast origin. Details of the approach are found in Polasek and Ren (2000). Also, we calculate the impulse response function for the conditional variances of the VAR-GARCH-M model using the following formula:

$$\Delta \hat{\mathbf{H}}_{t+s}(\varepsilon_{t+1}) = \frac{1}{s}[E_t(\mathbf{H}_{t+s}|\varepsilon_{t+1}, \Omega_t) - E_t(\mathbf{H}_{t+s}|\Omega_t)], \quad s = 1, 2, \ldots \tag{21}$$

4.6 Variance decomposition for the VAR-GARCH-M model

Given the MCMC output of size L, $\left\{\theta^{(l)}, l = 1, \ldots, L\right\}$ the mean, variances and covariances of the (e.g. one step ahead) forecasts for the VAR-GARCH-M model are given for all $j, m = 1, \ldots, M$ by

$$Ave(y^j_{t+1}) = \frac{1}{L}\sum_{l=1}^{L} y^{j(l)}_{t+1} = \bar{y}^j_{t+1}, \tag{22}$$

$$Var(y^j_{t+1}) = \frac{1}{L}\sum_{l=1}^{L} (y^{j(l)}_{t+1} - \bar{y}^j_{t+1})^2, \tag{23}$$

and

$$Cov(y^j_{t+1}, y^m_{t+1}) = \frac{1}{L}\sum_{r=1}^{L} (y^{j(l)}_{t+1} - \bar{y}^j_{t+1})(y^{m(l)}_{t+1} - \bar{y}^m_{t+1})'. \tag{24}$$

Collecting the variances $Var(y^j_{t+1})$ in (23) and the covariances in (24) into a covariance matrix $\mathbf{\Sigma}_{t+1}$ we can calculate the contribution of the innovations to the j^{th}-variance of the one-step ahead forecast. This procedure can be generalized so as to compute the proportion of the $j^{th}-$variance due to each innovation for a one-step ahead forecast.
For a VAR model the impulse response function (IRF) is obtained from the moving average representation (see Hamilton 1994)

$$\mathbf{y}_t = \mu + \varepsilon_t + \mathbf{M}_1\varepsilon_{t-1} + \mathbf{M}_2\varepsilon_{t-2} + \ldots \tag{25}$$

The impulse responses are the elements of $\mathbf{M}_s$ matrices $\{m_{ij,s}, i, j = 1, \ldots, M\}$ and are interpreted in the following way. If the residual ε_{t-s} is increased by one unit in the j^{th} component, then the value of the i^{th} variable $y_{i,t+s}$ at time $t+s$ is increased by $m_{ij,s}$ (holding all other influences constant). The value of the vector $\mathbf{y}_{t+s}$ at time $t+s$ of this simulation corresponds to the j^{th} column of the matrix $\mathbf{M}_s$. The whole matrix $\mathbf{M}_s$ is obtained by running ε_t in (25) from $\mathbf{e}_1$ to $\mathbf{e}_M$ over all unit vectors.
Since the impulse response function for the VAR-GARCH-M model cannot be given in closed form we obtain the IRF numerically in analogy to the VAR model. The variance decomposition is obtained in the following way:
1) We calculate the $\mathbf{M}_s$ matrices by choosing $\varepsilon_{t+1} = \mathbf{e}_j$ in the VAR-GARCH-M model and calculate the forecasts $\mathbf{y}_{t+s}$.

2) We calculate the covariance matrix $\mathbf{\Sigma}_s = Var(\mathbf{y}_{t+s})$ from the s-step ahead predictions in (17) or in (22)-(24) and we obtain from the Choleski decomposition the lower triangular matrix $\mathbf{P}_s$, i.e. $\mathbf{\Sigma}_s = \mathbf{P}_s\mathbf{P}_s'$.
3) For the s-step horizon $s = 1, \ldots, s_{max}$ we calculate the orthogonalized impulse responses $\mathbf{A}_{(s)} = \mathbf{M}_s\mathbf{P}_s$.
4) For the first equation we calculate, for $s = 1$, the relative variance decompositions. Denote the first row of $\mathbf{A}_{(s)}$ by $\underset{\sim}{\mathbf{a}}_1' = (a_{11}, \ldots, a_{1M})$ and the sum of squares is given by $s_{11}^2 = \underset{\sim}{\mathbf{a}}_1'\underset{\sim}{\mathbf{a}}_1$. The relative variance decomposition are given in the vector $\left(a_{11}^2, \ldots, a_{1M}^2\right)/s_{11}^2$.
For $s = 2$ we need the matrices $\mathbf{A}_{(1)}$ and $\mathbf{A}_{(2)}$ and for the first equation we need the first rows $\mathbf{a}_1^{(1)}$ and $\mathbf{a}_1^{(2)}$. The sum of squares is $s_{12}^2 = \mathbf{a}_1^{(1)'}\mathbf{a}_1^{(1)} + \mathbf{a}_1^{(2)'}\mathbf{a}_1^{(2)}$. The relative variance decomposition now is given by

$$\left(a_{11(1)}^2 + a_{11(2)}^2, \ldots, a_{1M(1)}^2 + a_{1M(2)}^2\right)/s_{12}^2.$$

In this way we can proceed to the s^{th} equation obtaining

$$\sum_{i=1}^{s}\left(a_{11(i)}^2, \ldots, a_{1M(i)}^2\right)/s_{1s}^2$$

with

$$s_{1s}^2 = \sum_{i=1}^{s}\sum_{m=1}^{M} a_{1m(i)}^2, \quad s = 1, \ldots, s_{max}.$$

For the other equations we proceed in the same way as for the first equation.

5 Empirical results for the US labour market

In this section we describe the data base and we discuss the main results obtained from estimating the five-dimensional VAR-GARCH-M model for the US labor market.

5.1 The data base

The time series for the empirical analysis are total employment and the employment shares of the durable, non-durable, transport and manufacturing sectors so that we have a five-dimensional VAR. We shall let n_t^j denote the natural logarithm of aggregate employment when $j = 1$ and instead indicate the logarithms of employment shares in the durable, non-durable, transport and service sectors respectively when $j = 2, 3, 4, 5$. We use quarterly growth rates $y_t^j = \triangle n_t^j = n_t^j - n_{t-4}^j = \log(N_t^j) - \log(N_{t-4}^j)$ of the United States quarterly employment data (1975Q1 - 1990Q4) from the OECD data base. This is the same data set used by MPZ (1995) for their experiment using dispersion measures. We resort to the same sample for aggregate and sectoral employment so as to check to see if our empirical analysis will confirm the support for the sectoral shift hypothesis that emerged from that work. By applying the same data set to a different model to investigate the same phenomenon, we subscribe to the methodological argument in Hendry, Leamer and Poirier (1990), according to which econometric modeling should be viewed as an incremental progressive accumulation of knowledge. The series 'Employment Civilian' is the total employment in our model (n_t), and we also take the Durable Goods Manufacturing employment (the sum of S24, S25, S32, S33, S34, S35, S36, S37, S38 and S39), the 'Non-durable Goods Manufacturing' employment (the sum of S20, S21, S22, S23, S26, S27, S28, S29, S30 and S31), the Transportation employment (the sum of S40, S41, S42, S44, S45, S46, S47), and the Services employment (the sum of SIC 70, 72, 73, 75, 76, 78-84, 86, 87 and 89) as the sector shares of employment.

5.2 Model selection

In the following we describe the results of the model selection procedure which has been carried out using the marginal likelihood criterion outlined in section 3.1. From Table 2 we can see the marginal likelihood, dependent on k (the order of the VAR(k) model), p and q the order of the GARCH model and r the number of lags in the ARCH-M component. The best model (which we shall call model 2 or M_2) is a VAR(2)-GARCH(2,2)-M(2) model with the largest marginal likelihood value of -244.50. We can compare the values in Table 2 with the marginal likelihoods of other models and perform a Bayes test.

Let us first test if the VAR(1) model (top row of Table 2) is worse than the best reported VAR-GARCH model. The first row of Table 2 shows a marginal likelihood of -253.77 and the best remaining model is still the VAR(2)-GARCH(2,2)-M(2) model the value -244.50 is marked with a star. Taking the difference yields $\log \mathrm{B}_{21} = 9.27$ ($\mathrm{B}_{21} \approx 10,615$). This means the VAR(1) model is 10,615 times less likely than the best VAR(2)-GARCH(2,2)-M(2) model. If we compare the value -248.45 of the best VAR(2)-GARCH(2,2)-M(0) model, i.e. a model where there is no feedback from the variances to the mean equation of the time series, then the difference in log marginal likelihoods is log $B_{21} = 3.95$, which yields a Bayes factor of $B_{12} = 0.0193$. This result implies that model 1 is about 51.94 times less likely than the best VAR-GARCH-M model. Thus the measure of relative support provided by the data for model 1 against model 2 is very strong against M1 in both cases.

5.3 How important are the GARCH components?

The model estimation procedure has been extended by imposing the following restrictions. We first consider the model where the first equation (aggregate employment) does not contain a GARCH component while the other sectoral equations are estimated with a GARCH structure (this is named restricted model 1: 'sectoral ARCH only'). Then we reverse the experiment by imposing a GARCH component only for the first equation (restricted model 2: 'total ARCH only') while the sectoral equations are simple VAR equations. The results are listed in the last two columns of Table 2. Subject to the first restriction, the best model is again a VAR(2)-GARCH(2,2)-M(2) model. Calling this model M2 and comparing it with the VAR(1) model, M1, we have a log $B_{21} = 10.64$, i.e. $B_{21} \approx 41,773$. If we compare M2 with the VAR(2)-GARCH(2,2)-M(0) or M1 model we compute the log-Bayes factor for M2 versus M1 to be 8.24 (i.e. $B_{21} \approx 3,789$). The empirical evidence for the US labour market strongly favours the model incorporating the GARCH effects even when these are confined only to the sectoral components.

If we consider the model where ARCH effects are only present in the total employment equation the best model is found to be a VAR(2)-ARCH(2,1)-M(2)model (Table 2). Comparing as before with the VAR(1) model we have a log $B_{21} = 8.18$ ($B_{21} \approx 3,569$). When we compare against the best model without the "M-component" we obtain a log-Bayes factor of 5.23 ($B_{21} \approx 187$). Thus, the GARCH-M model is again supported when only the aggregate shocks display a volatile structure. It should be pointed out that there is a drop in the value of the marginal likelihood of the best model when we move from the unrestricted model to the restricted ones. Note that the drop is larger for the model 'total ARCH only' where the GARCH component is limited only to the first equation. This result implies that there would be a bigger loss of information if we ignore the volatility effects of sectoral shocks than those of aggregate disturbances.

5.4 Variance decomposition analysis

In this section we present the results of the forecast error variance decomposition obtained as described in section 3.6. Sectoral shares are then introduced in the following order: durables, non-durables, services and transport. Since we are interested in having total employment ordered ahead of the sectoral shares, we have chosen one combination without attaching any particular economic interpretation to it. This triangularization implies a linear representation of a nonlinear model of shocks and needs cautious interpretation for the essence of sectoral shifts analysis but it can be interpreted as a 'linearized' lower bound on the non-linear contribution of sectoral shocks to the explanation of total employment variance ([11]).
Table 3 reports the results of the innovation accounting analysis carried out using the VAR(2)-GARCH(2,2)-M(2) model. The percentage of forecast error variance of aggregate employment growth accounted for by its own innovations and by innovations in the growth of sectoral shares in the first column are reported. If we consider a one-quarter horizon, the portion of the total variance in total employment due to reallocation shocks is approximately 30%. As s becomes larger the fraction of variance of total employment growth accounted for by sectoral reallocation shocks becomes larger. With a forecast horizon of four quarters, sectoral innovations contribute 65% to the aggregate employment variance. If we look at the forecast error variance of the sectoral components of the model, it should be noted that aggregate innovations cannot account for more than 28% of the sectoral variances whatever the forecast horizon. The one-step forecast error variance of sectoral components is largely (58-60%) accounted for by innovations within the own sector. However, as time is evolving the contributions of other sectors become larger. Thus, with a four-quarters horizon, other sectors innovations account for 53%, 34%, 41% and 47% of the variances of the durable, non-durable, transportation and services sectors, respectively.

5.5 The impulse response function

Finally we like to give a dynamic description of the VAR-GARCH-M model in terms of the impulse response functions of section 3.5. Table 4 shows the IRF for the last time point and the graph of the IRF is given in Figure 1. From the first panel we see that a shock in total employment has the highest effect after 4 periods on non-durable sectoral employment. The sectoral shocks seem to follow a common pattern. A positive shock in the first period is off-set by a negative shock of about the same size on period later. After 6 lags all IRF converge to zero.
Figure 2 shows the impulse response function for the conditional volatilities (the main diagonal) in $\mathbf{H}_{t+s}$ which are given in (21) if ε_{t+1} varies over all $M = 5$ unity vectors. Note that the volatilities of total employment and employment in the durable goods sector react most sensibly to shocks in any sector. Shocks in total employment does not seem to affect volatilities in the sectors: This means that aggregate shocks don't influence employment reallocations in the sectors. Except for the transportation sector, all sectoral shocks affects the volatility of the total employment growth rate. Employment reallocations in the durable goods sector seem to be highly exposed to growth shocks in other sectors.

6 Summary and concluding remarks

Extending previous sectoral employment growth models, we have estimated a VAR-GARCH-M model to incorporate the potential non-linearity of sectoral shocks. We used the MCMC algorithm to estimated the VAR-GARCH-M model that leads to an exact small sample distribution of the coefficients. For model selection we employed Bayes factors and the marginal likelihood criterion. We find that the ARCH structure for the aggregate and sectoral component is highly "significant". This results shows the importance of the volatility component for employment growth models and the necessity of modeling the implicit non-linearity of sectoral shifts models.
Two major results have emerged. First, when comparing hypotheses, Bayes factors suggest that the data support the VAR(2)-GARCH(2,2)-M(2) model relative to the other models. This result indicates that both volatility clustering and the feedback of volatilities into the growth rates play an im-

portant role. Second, the innovation accounting analysis, carried out using a Cholesky decomposition, where aggregate shocks are ordered ahead of sectoral innovations, shows that sectoral shocks account for approximately 65rate when we consider a one year forecast horizon. Thus reallocative shocks have a large and significant role in explaining total employment behaviour.

Our results are in agreement with the analysis of sectoral dynamics (and shocks) on aggregate unemployment in the models of Mills, Pelloni and Zervoyianni (1995), henceforth MPZ. MPZ is the only paper of the dispersion proxy generation providing strong evidence in favour of sectoral shifts, even when Lilien's dispersion measure is replaced by a variant "purged" of aggregate influences. They find that 55% of the variance of aggregate component growth is explained by the natural rate computed using their purged dispersion measure. The significant and large role MPZ find for sectoral shifts is highly consistent with our findings which have, however, the advantage of being derived directly from time series properties without having to appeal to ad hoc dispersion measures.

If we compare our findings with those of CK - where VAR models are introduced without dispersion proxies - we find that our VAR-GARCH-M model reveal more cross-sectoral dynamical results. Taking into account the fact they used a different sample period, different sectoral decompositions and a different VAR modelling strategy, we find that the VAR-GARCH-M model provides a stronger support to sectoral shifts than a VAR model without GARCH components. The CK's results vary quite widely according to the employed identifying restriction. When CK considers estimates of the reallocation shocks, the sectoral employment shares account for only 6% (bivariate VAR) and 27% (7-dimensional VAR) of the variance of total employment growth over a 2-year horizon and a 1-year horizon respectively. Instead over a 1-year horizon, as pointed out above, our findings suggest that a much higher proportion (65%) of the aggregate component variance can be explained by sectoral reallocations. We believe that this outcome is probably due to the fact that the class of VAR-GARCH-M models provides a new framework of sectoral employment growth models. We conclude that GARCH effects are important to explain the dynamics of sectoral employment growth models and throws new light on the impact of reallocation shocks.

Finally we wish to point out that our results are also related with Davis and Haltiwanger (1999, table 2, p. 1243) which reports positive serial correlation in the cross-sectional variance of idiosyncratic shocks ([12]). It seems that the ARCH components are important characteristics for sectoral employment models and adds a new dimension on the analysis of the aggregate effects of reallocation shocks: Once the heteroskedasticity of sectoral shocks is explicitly modelled, the volatility of sectoral shocks surfaces as an important source of non-linearity for sectoral labour market models. Also, we find additional evidence for this result in a separate paper dealing with major European economies (Pelloni and Polasek, 2000).

Notes

(1) All the models are estimated with the BASEL software package (see Polasek 2000).
(2) A third potential contribution of this paper would emerge if it were interpreted as a test of "certainty equivalence" as well. In models displaying certainty equivalence, heteroskedasticity in the driving processes should not affect agents' decision rules. However, it is reasonable to conceive that certainty equivalence breaks down for firms operating under fixed employment adjustment costs or linear hiring - firing costs (Caballero, Engel and Haltiwanger, 1997; Campbell and Fisher, 1996). If our paper is viewed in this perspective then current and lagged variance terms should not "significantly" enter the employment growth equations if certainty equivalence holds.
(3) Lilien's dispersion index can be written in our notation as

$$\hat{\sigma}_t = \left[\sum_{i=1}^{M} (N_{jt}/N_t) (\triangle \log N_{jt} - \triangle \log N_t)^2 \right]^{1/2}$$

where N_{jt} is employment in sector j and N_t is aggregate employment. Lilien then estimates a reduced form equation of the general form,

$$u_t = a + \sum_{i=1}^{T} b_i \sigma_{t-i} + \sum_{i=1}^{S} c_i x_{t-i} + \varepsilon_t$$

where x is a vector of aggregate shocks and u is the unemployment rate.
(4) Long and Plosser (1987) can be seen as a forerunner of this approach.
(5) CK on p.95 recognize how their symmetric treatment of sectoral shocks somehow departs from the standard view of sectoral shifts.
(6) For a discussion of Bayes factors c.f. Kass and Raftery (1995) and Poirier (1995). For Bayes factors and non-linear models see Koop and Potter (1997).
(7) Partial BF, local BF, pseudo BF, intrinsic BF, posterior BF, fractional BF, are the most recurrent concepts in the literature, c.f. Gelfand and Dey (1994), Kass and Raftery (1995), O'Hagan (1995) and the references therein.
(8) Let us have a sample y of size n and a training sample of size n_1, where both n and n_1 are large, then for a a given fixed proportion $b = n_1/n$, it follows that $f(x|\theta_j, M_j) \approx [f(y|\theta_j, M_j)]^b$ and we can calculate the fractional marginal likelihood

$$f(b, y|M_j) = \frac{\int f(y|\theta_j, M_j) f(\theta_j|M_j) d\theta_j}{\int [f(y|\theta_j, M_j)]^b f(\theta_j|M_j) d\theta_j}$$

(9) See Phillips (1991) with discussion and Uhlig (1994).
(10) Since unit root tests can be affected by structural breaks and outliers we run a Bayes test for a possible break point in the time series. We can see from Table 1 that the stationary (DF) outlier models always perform better than the non-stationary models. All outlier models are slightly better than the original models in Table 1. This outcome suggests that the step changes observed in the beginning of the 80's are rather outlier effects than structural breaks; c.f. Pelloni and Polasek (1999) for further details.
(11) CK adopt a similar device in order to implement their analysis.

Order	Total %			
p	$\Delta - AR$	$DF - AR$ 1	$DF - AR$ 2	$DF - AR$ 3
1	190.1606	193.2524	193.2577	192.7290
2	191.4274*	195.4222*	197.7968**	194.3029*
3	191.2504	194.6288	197.1488	193.6504
Order	Durable goods(share %)			
p	$\Delta - AR$	$DF - AR$ 1	$DF - AR$ 2	$DF - AR$ 3
1	148.8612	153.4889	154.1387	151.0513
2	149.2258	153.7205	155.9291	153.2917
3	149.5605	154.4872	160.2191	158.5116
4	161.9327	164.5190	165.4652	162.5935
5	163.2937*	165.3017*	165.8435**	162.7422*
6	162.6695	164.7458	165.7810	162.7358
Order	Non-durable goods(share %)			
p	$\Delta - AR$	$DF - AR$ 1	$DF - AR$ 2	$DF - AR$ 3
1	178.3648	180.7102	185.5339	181.9984
2	178.5722	180.6508	184.6048	181.1417
3	182.8687	185.4732	190.9893	187.6614
4	185.2309*	187.2895*	191.2793**	187.8355*
5	185.2048	187.1330	191.0991	187.6716
Order	Transportation(share %)			
p	$\Delta - AR$	$DF - AR$ 1	$DF - AR$ 2	$DF - AR$ 3
1	187.0468	191.5356	189.8160	186.3851
2	187.7452	193.3442	191.9324	188.7775
3	189.2631	196.9206	196.3297	195.6773
4	201.4560*	206.1169**	204.4229*	201.6194*
5	200.7137	205.2555	203.5438	200.7393
Order	Services(share %)			
p	$\Delta - AR$	$DF - AR$ 1	$DF - AR$ 2	$DF - AR$ 3
1	194.2419	197.6030	199.4266	195.8750
2	195.0135	197.7052	201.7799	198.2570
3	195.4949	198.5406	204.1732	200.6853
4	208.4259*	210.4122*	212.5920**	208.9598*
5	208.1274	210.0818	211.0447	208.2900

Table 1. Bayesian unit root tests: The fractional log marginal likelihood of US total employment and employment shares for augmented Dickey-Fuller AR(p) models from 1975 Q1 to 1990 Q4.

order				full model	'sectoral ARCH only'	'total ARCH only'
k	p	q	r		restricted model 1	restricted model 2
1	0	0	0	-253.77	-261.43	-263.54
1	1	1	1	-246.22	-253.91	-257.33
1	2	3	1	-244.72	-252.17	-258.76
2	1	1	2	-245.33	-253.09	-257.13
2	1	2	1	-245.53	-255.08	-257.58
2	1	3	2	-246.75	-258.44	-260.08
2	2	1	2	-247.22	-256.91	-255.36^*
2	2	2	0	-248.45	-259.03	-260.59
2	2	2	1	-245.76	-256.90	-256.56
2	2	2	2	-244.50^*	-250.79^*	-256.48
2	3	3	2	-247.89	-254.09	-257.68
3	3	3	3	-246.11	-255.12	-258.90
3	3	3	2	-246.10	-257.87	-260.08
3	1	1	2	-245.45	-258.43	-261.2
3	1	1	3	-245.09	-257.33	-260.31
3	1	2	3	-248.81	-258.08	-260.91
3	1	2	2	-249.87	-259.09	-260.88
3	1	3	2	-246.54	-260.32	-261.04
3	1	3	3	-249.78	-260.55	-261.68

Table 2. The log marginal likelihoods for VAR(k)-GARCH(p,q)-M(r) models: The full and the restricted models

Time	Variance decomposition in				
Quarter	Total	Durables	Nondur-ables	Transpor-tation	Services
1	0.6959	0.2206	0.0361	0.1849	0.1685
	0.0459	0.5876	0.2634	0.0297	0.0556
	0.1609	0.1237	0.6015	0.1683	0.1391
	0.0852	0.0532	0.0881	0.6085	0.0572
	0.0121	0.0149	0.0109	0.0085	0.5795
2	0.5165	0.1678	0.0011	0.1824	0.1865
	0.1090	0.4620	0.1047	0.1066	0.0347
	0.2482	0.2431	0.4534	0.2599	0.0653
	0.1246	0.1257	0.1320	0.4480	0.0911
	0.0018	0.0014	0.3088	0.0031	0.6223
3	0.3835	0.1884	0.2573	0.2781	0.1780
	0.1183	0.2894	0.1297	0.1314	0.1195
	0.2027	0.2074	0.3099	0.2046	0.1785
	0.2151	0.2335	0.2273	0.2960	0.2349
	0.0804	0.0812	0.0759	0.0899	0.2891
4	0.3485	0.2166	0.2442	0.2621	0.2401
	0.0276	0.2578	0.0276	0.0316	0.0197
	0.2916	0.2389	0.4160	0.2848	0.2102
	0.2417	0.2285	0.2199	0.3246	0.2371
	0.0905	0.0583	0.0924	0.0969	0.2929

Table 3. Variance decomposition of US employment for the VAR(2)-GARCH(2,2)-M(2) model in percentages

		Impulse response in				
Data	Quarter	Total	Durable	Non-durable	Transpor.	Services
Total	1	1.0000	0.0000	0.0000	0.0000	0.0000
	2	0.3459	1.2514	1.2995	-0.4266	-0.1264
	3	0.4819	0.9174	1.1538	-0.4376	-0.0734
	4	0.6844	1.0238	1.9930	0.0560	0.3424
	5	0.6908	0.7873	1.8992	0.4505	0.5052
	6	0.2233	0.0244	0.2792	0.2171	0.0828
	7	0.0689	0.0220	-0.1198	-0.0480	-0.0774
	8	0.0889	0.0788	-0.1078	-0.0743	-0.1036
Durable	1	0.0000	1.0000	0.0000	0.0000	0.0000
	2	0.5863	-0.1579	-0.9491	-0.5665	-1.1761
	3	0.2578	-0.0077	-0.9702	-0.6539	-1.0178
	4	0.1650	0.1559	-0.4258	-0.3310	-0.6754
	5	0.0718	0.0823	-0.3165	-0.0844	-0.4075
	6	0.1139	0.1235	-0.0519	0.0072	-0.0995
	7	0.0978	0.0601	-0.1215	-0.0176	-0.0732
	8	0.0912	0.0795	-0.1089	-0.0708	-0.1032
Nondur.	1	0.0000	0.0000	1.0000	0.0000	0.0000
	2	0.5212	0.4909	-1.8551	-0.6325	-1.1948
	3	0.2058	0.1928	-1.3640	-0.7050	-1.0317
	4	0.1251	0.1990	-0.5933	-0.3696	-0.6797
	5	0.0678	0.0961	-0.2944	-0.0917	-0.3821
	6	0.1414	0.1504	0.0484	0.0326	-0.0574
	7	0.0896	0.0437	-0.1411	-0.0216	-0.0799
	8	0.0902	0.0779	-0.1115	-0.0714	-0.1041
Transpor.	1	0.0000	0.0000	0.0000	1.0000	0.0000
	2	0.8842	1.2394	-0.0113	-1.1573	-0.9178
	3	0.5368	0.8449	-0.1246	-0.7806	-0.7674
	4	0.4742	0.8365	0.4540	-0.2502	-0.3891
	5	0.3302	0.5210	0.3300	0.0752	-0.1739
	6	0.1102	0.1028	-0.1668	0.0214	-0.1354
	7	0.1011	0.0865	-0.1003	-0.0271	-0.0700
	8	0.0920	0.0827	-0.1047	-0.0714	-0.1021
Services	1	0.0000	0.0000	0.0000	0.0000	1.0000
	2	0.7703	1.0730	-0.7323	-0.5985	-1.8189
	3	0.4126	0.6732	-0.7835	-0.6654	-1.2155
	4	0.3078	0.6240	-0.2606	-0.3210	-0.7164
	5	0.1554	0.3326	-0.2820	-0.0708	-0.4238
	6	0.0880	0.1383	-0.2274	-0.0291	-0.1726
	7	0.1067	0.0918	-0.1064	-0.0196	-0.0707
	8	0.0923	0.0824	-0.1058	-0.0705	-0.1022

Table 4. Individual impulse response function for US (total and sectoral) employment for the VAR(2)-GARCH(2,2)-M(2) model

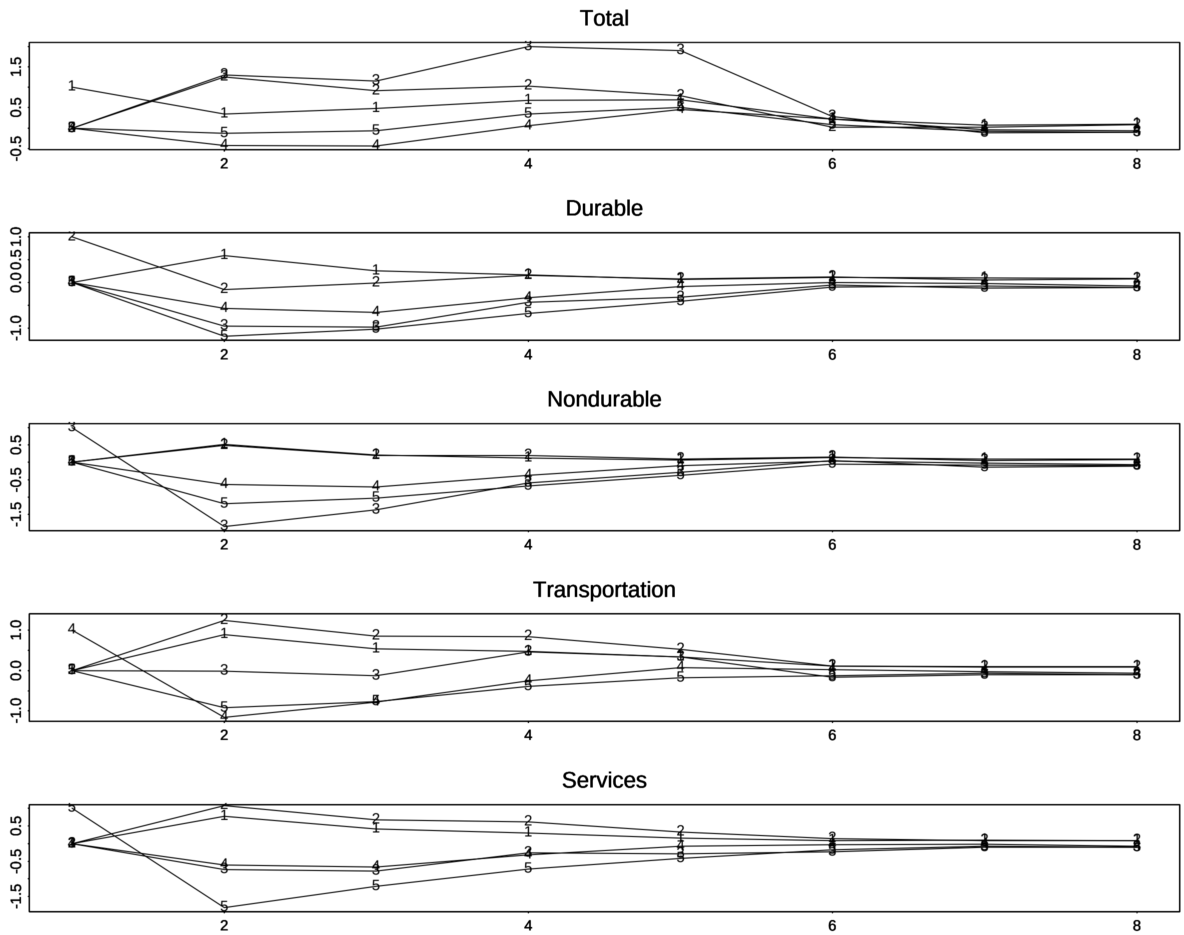

Figure 1: Individual impulse response plots of US employment (for the growth rates) from Q1 1975 to Q4 1990 for the VAR(2)-GARCH(2,2)-M(2) model, where 1=Total, 2=Durables, 3=Non-durables, 4=Transportation and 5=Services

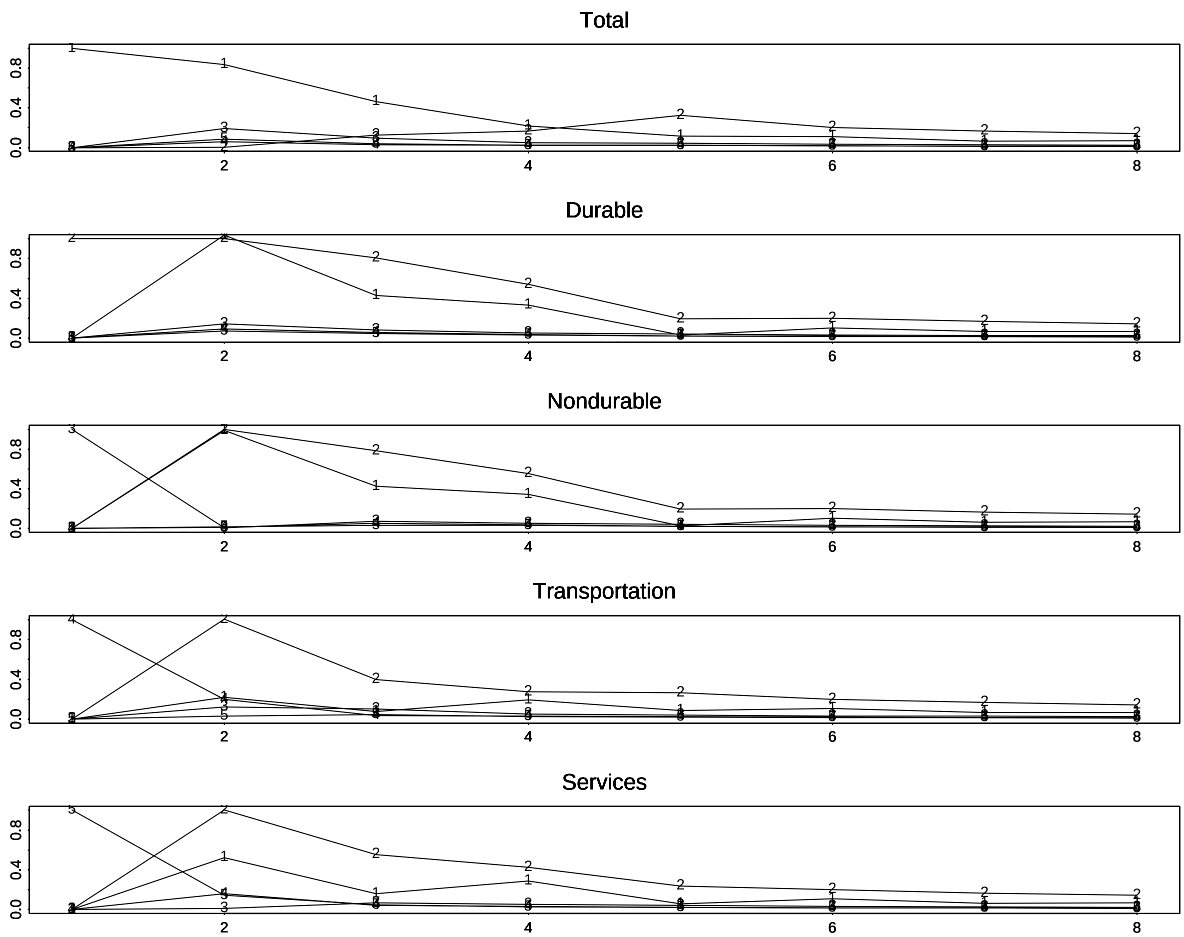

Figure 2: Individual impulse response plots of US employment (for the volatilities) from Q1 1975 to Q4 1990 for the VAR(2)-GARCH(2,2)-M(2) model, where $1 = Total$, $2 = Durables$, $3 = Non-durables$, $4 = Transportation$ and $5 = Services$

Acknowledgments

We wish to thank John Addison, Tom Crossley, Giovanni Gallipoli, Jurek Konieczny, Gary Koop, participants in the CEA99 meeting (Toronto, May 1999), the IEFS-UK 99 meeting (London, April 1999), the departmental seminars at York University (CND) and Wilfrid Laurier University and the workshop "Dynamics of Deliberation, Probability and Economics" at Trento University (October, 1997) for helpful comments and suggestions. We thank Lei Ren and Song Jin for the programming help. The usual disclaimer applies.

$*$ Financial support from the National Research Council of Italy (CNR) under the projects 96.01649.CT10, 97.01378CT10 and 97.01395.CT10 is gratefully acknowledged.

$**$ Financial support from the National Research Council of Italy (CNR) under the project 96.01649.CT10 and 97.01378CT10 is gratefully acknowledged.

References:

Abraham, K. and Katz, L.F. (1984): Cyclical unemployment: Sectoral Shifts or Aggregate Disturbances?. *National Bureau of Economic Research, working paper no. 1410, July.*

Abraham, K. and Katz, L. (1986): Cyclical Unemployment: Sectoral Shifts on Aggregate Disturbances. *Journal of Political Economy, 94, 507-522.*

Aitkin, M. (1991): Posterior Bayes Factors, (with discussion), *Journal of the Royal Statistical Society, B, 53, 111-142*

Bollerslev, T., Engle, R.F. and Wooldridge, J.M. (1988): A capital-asset pricing model with time-varying covariances. *Journal of Political Economy, 96, 116-31.*

Bollerslev, T., Chou, R.U. and Kroner, K.F. (1992): ARCH modelling in finance: a review of the theory and empirical evidence. *Journal of Econometrics, 52, 5-59.*

Brainard, S.L. and Cutler, D.M. (1993): Sectoral shifts and cyclical unemployment reconsidered. *Quarterly Journal of Economics, 108, 219-243.*

Caballero, R.C., Engel, E.M.R.A. and Haltiwanger, J. (1997): Aggregate employment dynamics: Building from microeconomic evidence, *American Economic Review, 87, 115-137.*

Campbell, J.R. and Fisher, D.M. (1996): Aggregate employment fluctuations with microeconomic asymmetries, *NBER, WP5767, September.*

Campbell, J.R., Kuttner, K.N. (1996): Macroeconomic effects of employment reallocation. *Carnegie-Rochester Conference Series on Public Policy, 44, 87-116.*

Campbell, J.Y., Lo, A.W. and MacKinlay, A.C. (1997): *The Econometrics of*

1234-1258.

Dey, D.K. (1994): Bayesian Model Choice: Asymptotics and Exact Calculations, *Journal of the Royal Statistical Society, B, 56, 501-514.*

Engle, R.F. (1982): Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica, 50, 987-1008.*

Engle, R.F. (1983): Estimates of the variance of U.S. inflation based upon the ARCH model. *Journal of Money Credit and Banking, 15, 286-301.*

Engle, R.F.and Kroner, K.F. (1995): Multivariate Simultaneous Generalized ARCH. *Econometric Theory, 11, 122-150.*

Gallipoli, G. and Pelloni, G. (2000): Macroeconomic effects of employment reallocation: a review and an appraisal from an econometric perspective, Discussion Paper 2000-10, Department of Economics, *mimeo,* University of Sheffield.

Gordon, S. and Boccanfuso, D. (1998): What can we learn from structural vector autoregression models?, *mimeo,* University of Laval, November.

Hamilton, J. (1994): *Time Series Analysis.* Princeton University Press, New York.

Hendry, D.F., Leamer, E.E. and Poirier, D.J. (1990): The ET dialogue: A conversation on econometric methodology, *Econometric Theory, 171-261.*

Kass, R.E., and Raftery A.E. (1995): Bayes Factors, *Journal of the American Statistical Association, 90, 773-795.*

Koop, G., Pesaran, H. and Potter, S.M. (1996): Impulse response analysis in nonlinear multivariate models, *Journal of Econometrics, 74, 119-148.*

Koop, G. and Potter, S.M. (1999): Bayes factors and nonlinearity: evidence from economic time series, *Journal of Econometrics, 88, 251-281.*

Lilien, D.M. (1982a): Sectoral shifts and cyclical unemployment. *Journal of Political Economy, 90, 777-793.*

Lilien, D.M. (1982b): A sectoral model of the business cycle, *MRG Working Paper no. 8231,* University of Southern California.

Liu S. and Polasek W. (1999): Maximum Likelihood Estimation for the VAR-VARCH Model: A New Approach, *in: U. Leopold-Wildburger, G. Feichtinger and K.-P. Kistner (eds.) Modelling and Decisions in Economics, Physica Verlag, Heidelberg, 99-114.*

Litterman, R.B. (1986): A statistical approach to economic forecasting. *Journal of Business and Economic Statistics, 4, 1-24.*

Long, J.B and Plosser, C.I. (1987): Sectoral vs. aggregate shocks in the business cycle. *American Economic Review, 77, 333-336.*

Loungani, P. and Rogerson, R. (1989): Cyclical fluctuations and the sectoral reallocation of labor: evidence from the PSID. *Journal of Monetary Economics, 23, 259-273.*

Loungani, P., Rush, M. and Tave, W. (1990): Stock market dispersion and unemployment. *Journal of Monetary Economics, 25, 367-388.*

Lucas, R.E. Jr. (1973): Some international evidence on output-inflation trade-offs. *American Economic Review, 63, 326-334.*

Lucas, R.E. Jr. and Prescott, E.C. (1974): Equilibrium search and unemployment. *Journal of Economic Theory, 7, 188-209.*

Mills, T.C., Pelloni, G. and Zervoyianni, A. (1995): Unemployment fluctuations in the United States: further tests of the sectoral shifts hypothesis. *The Review of Economics and Statistics, 77, 294-304.*

Mills, T.C., Pelloni, G. and Zervoyianni, A. (1996): Cyclical unemployment and sectoral shifts: further tests of the Lilien hypothesis for the U.K. *Economics Letters, 52,*

Murphy, K.M. and Topel, R.H. (1987): The evolution of unemployment in the United States, NBER. *Macroeconomics Annual, 2, 11-58.*

Neuman, G.R. and Topel, R.H. (1991): Employment risk, diversification, and unemployment. *Quarterly Journal of Economics, 106, 1341-1366.*

O'Hagan, A. (1995): Fractional Bayes factors for model comparison, (with discussion), *Journal of the Royal Statistical Society, Soc. B, 57, 99-138.*

Pelloni, G. and Polasek, W. (1999): Intersectoral Labour Reallocation and Employment Volatility: A Bayesian analysis using a VAR-GARCH-M model, *Discussion Paper no.99/4, Department of Economics, University of York (UK).*

Pelloni, G. and Polasek, W. (2000): Macroeconomic effects of sectoral shocks in U.S., U.K. and Germany: A BVAR-GARCH-M approach, *WP 200-03, Department of Economics, University of Sheffield.*

Phillips, P.C.B. (1991): To criticize the critics: an objective Bayesian analysis of stochastic trends, (with discussion), *Journal of Applied Econometrics, 6, 333-364.*

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

Polasek, W. (2000): *The BASEL package,* PC-Version 1.0. *University of Basel* *http://www.unibas.ch/iso/.*

Polasek, W. and Ren, L. (1999): Linear model selection in the presence of outliers and break points, *Institute of Statistics and Econometrics, University of Basel* *http://www.unibas.ch/iso/.*

Polasek, W. and Ren, L. (2000a) Generalized Impulse Response Functions for VAR-GARCH-M Models, *in: W. Gaul et al. (eds.) Data Analysis, Scientific Modeling and Practical Application p.299-311, Springer Verlag 2000 .*

Polasek, W. and Ren, L. (2000b) Volatility Analysis During the Asia Crisis: A multivariate GARCH-M model for stock returns in the US, Germany and Japan, *to appear in: Applied Stochastic Models in Bus. and Industry.*

Starr-McCluer, M. (1993): Cyclical fluctuations and sectoral reallocation: a re-examination. *Journal of Monetary Economics, 31, 417-425.*

Uhlig, H. (1994): What macroeconomists should know about unit roots, *Econometric Theory, 10, 645-671.*

Appendix

The Bayesian VAR(k)-GARCH(p,q)-M(r) model is then given by

$$\mathbf{Y} \sim \mathcal{N}_{T\times M}[\mathbf{BX} + \mathbf{\Psi}\tilde{\mathbf{H}}, diag(\mathbf{H}_1, \ldots, \mathbf{H}_T)], \tag{26}$$

$$vech\mathbf{H}_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i vech(\mathbf{u}_{t-i}\mathbf{u'}_{t-i}) + \sum_{j=1}^{p} \mathbf{\Phi}_j vech\mathbf{H}_{t-j},$$

and the prior distributions are chosen from the families of normal distributions, hence

$$\mathbf{B} \sim \mathcal{N}_{M\times(1+\tilde{M}k)}[\mathbf{B}_*, \mathbf{\Sigma}_{B_*} \otimes \mathbf{I}_M],$$

$$\mathbf{\Psi} \sim \mathcal{N}_{M\times \tilde{M}r}[\mathbf{\Psi}_*, \mathbf{\Sigma}_{\Psi_*} \otimes \mathbf{I}_M],$$

where all of the hyper-parameters (which are denoted with a star) are known a priori.
If $\mathbf{H} = diag(\mathbf{H}_1, \ldots, \mathbf{H}_T)$ is a $TM \times TM$, $\mathbf{W}$ a $r \times T$, and $\mathbf{V}$ a $T \times k$ matrix, then we define the special matrix

$$\begin{aligned} < \mathbf{w}'_t \mathbf{H}_t \mathbf{v}_t >_{rM\times kM} &= (\mathbf{W} \otimes \mathbf{I}_M) diag(\mathbf{H}_1, \ldots, \mathbf{H}_T)(\mathbf{V} \otimes \mathbf{I}_M) \\ &= \begin{pmatrix} \sum_t w_{1t}\mathbf{H}_t v_{t1}, \ldots, \sum_t w_{1t}\mathbf{H}_t v_{tk} \\ \ldots \\ \sum_t w_{rt}\mathbf{H}_t v_{t1}, \ldots, \sum_t w_{rt}\mathbf{H}_t v_{tk} \end{pmatrix}. \end{aligned}$$

For the Metropolis-Hastings step we use a normal distribution as proposal density

$$q(\alpha^{(j-1)}, \alpha^{(j)}) = \mathcal{N}[\alpha^{(j-1)}, \tilde{\mathbf{V}}], \tag{27}$$

where $\tilde{\mathbf{V}} = diag(\tilde{\mathbf{\Sigma}}_{\alpha_0}, \tilde{\mathbf{\Sigma}}_{\alpha_1}, \ldots, \tilde{\mathbf{\Sigma}}_{\alpha_p}, \tilde{\mathbf{\Sigma}}_{\mathbf{\Phi}_1}, \ldots, \tilde{\mathbf{\Sigma}}_{\mathbf{\Phi}_p})$ is a block diagonal variance covariance matrix. $q(\alpha^{(j-1)}, \alpha^{(j)})$ denotes the proposal density for moving from $\alpha^{(j-1)}$ to $\alpha^{(j)}$ and the variance blocks are approprietely chosen
We calculate the marginal likelihood as in Chib an Jeliazkov (1999) for suitable chosen parameter values $\hat{\alpha}$ and $\hat{\mathbf{B}}$ as the ratio of two sampling averages:

$$\hat{p}(\hat{\alpha} \mid \mathbf{Y}) = \frac{\bar{m}}{\bar{m}_0} \tag{28}$$

$$\bar{m} = \frac{1}{G} \sum_{g=1}^{G} A(\alpha^{(g)}, \hat{\alpha} \mid \mathbf{Y}, \mathbf{B}^{(g)}) \cdot q(\alpha^{(g)}, \hat{\alpha} \mid \mathbf{Y}, \mathbf{B}^{(g)}), \tag{29}$$

where $(\hat{\alpha}^{(g)}, \hat{\mathbf{B}}^{(g)})$ comes form the total $MCMC$ sample and

$$\bar{m}_0 = \frac{1}{J}\sum_{j=1}^{J} A(\alpha^{(j)}, \hat{\alpha} \mid \mathbf{Y}, \mathbf{B}^{(j)}).$$

is calculated from the sample $\{\alpha^{(j)}, \mathbf{B}^{(j)} \mid j = 1, \ldots, J\}$ which comes from the reduced $MCMC$ run with the conditional distribution

$$p(\mathbf{B} \mid \mathbf{Y}, \hat{\alpha}) \quad = \quad \mathcal{N}[\mathbf{B}_{**}(\hat{\alpha}), \mathbf{H}_{**}(\hat{\alpha})],$$

and the moments

$$\mathbf{H}_{**}^{-1}(\hat{\alpha}) \quad = \quad \mathbf{I}_M \otimes \mathbf{\Sigma}_{B_*}^{-1} + < \mathbf{x}_t^{'} \mathbf{H}_t^{-1} \mathbf{x}_t >, \tag{30}$$

$$\hat{\mathbf{B}}_{**}(\hat{\alpha}) \quad = \quad \mathbf{H}_{**}[vec(\mathbf{\Sigma}_{B_*} \mathbf{B}_* + < \mathbf{x}_t^{'} \mathbf{H}_t^{-1} \mathbf{y}_t >)], \tag{31}$$

where the matrices are partioned as $\mathbf{Y} = (\mathbf{y}_1, \ldots, \mathbf{y}_T)'$ and $\mathbf{X} = (\mathbf{x}_1, \ldots, \mathbf{x}_T)$. The parameters $\alpha^{(j)}$ are drawn form the proposal distribution (27) and satisfy the stationarity condition of VARCH models. Finally, the log marginal likelihood is calculated as

$$\log\hat{\mathrm{p}}(\mathbf{Y}) = \log\mathrm{p}(\mathbf{Y} \mid \hat{\alpha}, \hat{\mathbf{B}}) + \log\mathrm{p}(\hat{\alpha}, \hat{\mathbf{B}}) - \log\hat{\mathrm{p}}(\hat{\alpha} \mid \mathbf{Y}) - \log\hat{\mathrm{p}}(\hat{\mathbf{B}} \mid \mathbf{Y}, \hat{\alpha})$$

where the first term is the log-likelihood ordinate, the second one is the prior ordinate and the posterior estimated densities are given in (28). The last ordinate is calculated as

$$\hat{p}(\hat{\mathbf{B}} \mid \mathbf{Y}, \hat{\alpha}) = \mathcal{N}[\hat{\mathbf{B}} \mid \mathbf{B}_{**}(\hat{\alpha}), \mathbf{H}_{**}(\hat{\alpha})]. \tag{32}$$

The acceptance probability in (28) is calculated as

$$A(\alpha^{(j)}, \hat{\alpha} \mid \mathbf{Y}, \mathbf{B}^{(j)}) = min\left(1, \frac{p(\hat{\alpha} \mid \mathbf{Y}, \mathbf{B}^{(j)})}{p(\alpha^{(j)} \mid \mathbf{Y}, \mathbf{B}^{(j)})}\right)$$

where the conditional posterior density $p(\alpha \mid \mathbf{Y}, \mathbf{B})$ is proportional to the multivariate density in (26).