

Abstract

The purpose of this study is to test the Brenner hypothesis using a Bayesian vector autoregressive approach, to compute impulse response function and to conclude if evidence of the hypothesis exist in Canadian data.

Contents

1	Introduction:	2
2	Review of literature:	3
3	Methodology and results:	3
	3.1 data :	4
	3.2 Simple data statistic :	5
	3.3 Cochrane variance ratio test :	5
	3.4 Ols and polynomial distributed lag models :	6
	3.5 Specification Using Bayesian Vector Autoregression (BVAR) :	8
	3.6 Impulse responses and variance decomposition :	11
4	Extensions and Conclusion :	15
5	References :	15

On Unemployment and Mortality in Canadian Data; The Brenner Hypothesis revisited.

Soderstrom, Lee* Mikhail, Ossama[†]

November 23, 2000

1. Introduction:

The paper explore the casual relation between mortality and unemployment (UE) by using time series techniques. The pioneering work of Brenner(1979) exposed such relationship and concluded evidence of it. In his seminal work, psychosocial stress is the channel through which unemployment have influence over mortality, which increase the social cost of higher unemployment due to a recession or to a consequence of public policies such as disinflation. Recently, there is a growing concern that the social cost of unemployment is underestimated by public policy makers. The Brenner hypothesis states that economic instability manifested in periods of high unemployment originate sufficiently stressful life changes such that the mortality experience of the population concerned will be affected. Mortality can be associated with recessions long after the period of initial economic downturn. As Gravelle et al (1981) summarized it ” economic trauma that initiate a chain of distressing events ”. The hypothesis argues that unemployment influence health for any or all of three reasons: (a) psychological factor resulting from stress, (b) risk factor resulting from increased consumption of goods that may induce health deterioration such as alcohol or tobacco, and (c) demand factor resulting from a significant decrease of consumption of necessary goods due to the cut in the financial budget of the household. If evidence of the hypothesis

*Department of Economics, McGill University.

[†]Graduate Student, Department of Economics.

is found, decision makers might understand more the social consequences of unemployment and may want reduce its effects. The purpose of this study is to test the Brenner hypothesis in Canadian data. The paper begins in section II with a review of the literature. Section III presents the methodology and the results. Section IV concludes and suggest extensions.

2. Review of literature:

Brenner (1979) tested his claim for England and Wales during the period of 1939-1976. The study concluded that the decline in mortality over that period was due to the reduction of the severity of economic fluctuations, specially post WWII. Several recent studies attempted to test the existence of the Brenner hypothesis using different data sets. Among others, the rate of first admission to mental hospital in New-York was found inversely related to employment during the period 1914-1967 and to the business cycle index during 1841-1909. Also, data on incidences of acute pathological disturbances (including suicides and mortality) was found to be positively correlated to the unemployment rates within a year period.

Critics to the Brenner study came at a technical level rather than at a theoretical one. Which gave the hypothesis its status among key decision makers and raised the level of awareness of public policies social costs. Such critics as the specification of the regression equation, omitted/included variables questions, reliability of the data and robustness of the results had created a need to investigate the Brenner hypothesis at technical level. Here, one attempt to apply existing methods to such aim.

3. Methodology and results:

Quantitative research in economics can be divided into two approaches : system of equations approach (SEA) and calibration approach (CA). Both differ in ; wether the model is designed to possess a steady state, the extent of information available to the agents, the specification of the dynamic structure of the model and the nature of the exogenous variables. Here, the point of interest in designing a model to test the Brenner hypothesis is the dynamic structure part. The former approach rely on a structure that is determined from the data, not driven by any formal optimization and is carried on by an unrestricted VAR model (a system of dynamic linear equations). The latter approach focus on theory to determine

the framework of the dynamics and is carried by a restricted VAR model. Here, an obvious choice to test the Brenner hypothesis is the use an unrestricted VAR, which rely on very little theory and is more flexible in the sense that the dynamic structure of the model is determined from the data. One might ask about the validity and the benefits of cross-section versus time series studies in testing the hypothesis. Economic data suffer from a very low signal-to-noise ratio, which leaves little confidence about the useful economic structure at hand. The former suffer from selection bias. This bias is pronounced in the following reasoning, healthier workers are more mobile and move to lower unemployment areas which attracts healthier job seekers. Such mobility effect will lower mortality rates in low unemployment areas. Yet, still one can attempt to correct for such selection bias, and that is not a deterrent to the cross-section approach. Our interest here is on the later framework as a first step in testing the Brenner hypothesis in Canadian data by using a vector autoregressive (VAR) approach. VAR models became very useful as a tool for economic analysis; to assess the in-sample additional predictive content of one variable for another, to assess the in-sample effect of a typical shock on the rest of the system by mean of using impulse response functions and variance decompositions, and finally to understand policy changes by out-of-sample unconditional and conditional forecasts to assess the effect of a policy change.

Many critics to VAR showed that seasonal adjustment induce noninvertible MA components in the adjusted data. The issue of variable ordering¹ is a major caveat of using VAR approach. Here, we use the ordering of the variables to our advantage. Since the death data does not influence unemployment, one simply place the later at the first position in the VAR so that the later adds no information to explaining it. On the issue that VAR is a linear system, note that non-linearity in the conditional mean is of marginal interest here, since monthly and quarterly macroeconomic time series are usually well approximated by linear processes².

3.1. data :

Canadian data³ post early 70s provides a rich ground for testing the hypothesis. In 1971 unemployment insurance reform and the medicare program provided good quality for health services, therefore a data evidence of casual relationship between

¹Formally discussed latter.

²Brock and Sayers (1988) "Is the Business Cycle Characterized by Deterministic Chaos ?" Journal of Monetary Economics, 22, pp 71-90.

³Death : CANSIM series D157. Unemployment : CANSIM series D980342.

unemployment and mortality - if found - could shed some light on the hypothesis. Here, the data are not detrended to avoid the "magic of detrending"⁴ noted in Kasl (1979). Also, we use regional data to avoid the "ecological fallacy"⁵ of Gravelle (1981).

3.2. Simple data statistic :

Simple statistics on the level of UE and mortality. Quarterly Data From 1976:01 To 1997:02.

	Mean	Std Error	Median
Unemployment	46541.25	4414.34	45654
Death	1244.11	287.62	1283

Cross Correlations

Lag	Cross Correlations					
-9:	0.368	0.375	0.289	0.336	0.476	0.498
-3:	0.410	0.464	0.621	0.649	0.541	0.542
+3:	0.640	0.630	0.502	0.507	0.613	0.620
+9:	0.510	0.506	0.598	0.584	0.445	0.407

3.3. Cochran variance ratio test :

The cochrane test serves to clarify that both series exhibit the same pattern of persistence. Early, the use of polynomial distributed lag models was criticized as "time series fallacy"⁶ methods. This test is to show that both time series display the same persistence pattern, which makes them primary objects in a casual relationship analysis. The cochrane test computes - for a general variable y_t - the following :

$$V^k = \frac{1}{k+1} \frac{\text{var}(y_{t+k+1} - y_t)}{\text{var}(y_{t+1} - y_t)} = 1 + 2 \sum_{j=1}^k \left(1 - \frac{j}{k+1}\right) \rho_{j(\Delta y)} \quad (3.1)$$

⁴Kasl questioned the detrending procedures and claimed no rationale for using it - in this context - since it usually result in loss of information.

⁵Using macro-level (aggregate, ecological) studies can produce high correlations between the variables which are likely to be reflecting spurious associations only.

⁶When two or more time series are said to be related, when in fact the statistical association reflects a casual covariance due only to concomitant time series structure in the variation pattern of individual series.

where $\rho_{j(\Delta y)}$ is the j^{th} order autocorrelation coefficient for the first-difference of y_t and k is the lag. If the variable is stationary, then V^k tend to zero as $k \rightarrow \infty$.

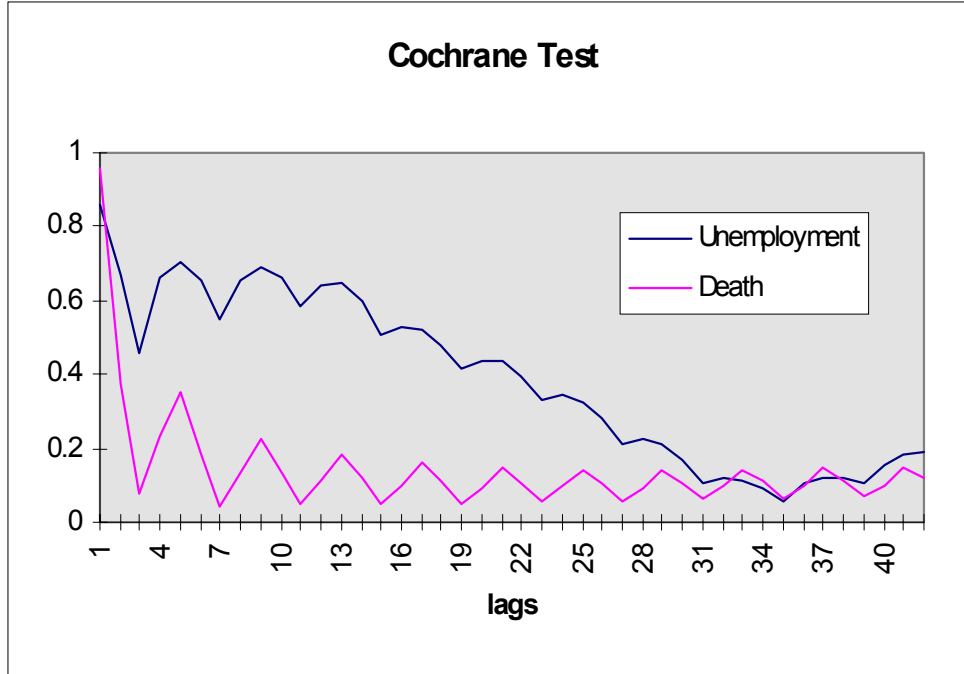


Figure 3.1:

3.4. Ols and polynomial distributed lag models :

In this section a simple OLS is performed and a replication of the Brenner model is conducted⁷. Briefly, the distributed lag model used at the 2nd degree polynomial is reproduced. The regression :

$$m_t = const + \sum_{i=0}^T \beta_i \cdot U E_{t-i} + e_t \quad (3.2)$$

where the β 's are restricted to behave as a 2nd degree polynomial, i.e.

$$\beta_i = \lambda_0 + \lambda_1 \cdot i + \lambda_2 \cdot i^2 \quad (3.3)$$

⁷We replicate the model using the mortality and unemployment series. Brenner's original study included variables to account for economic conditions; such as government transfer payments as a proportion of total expenditures and real per capita personal disposable income.

with no end points restrictions. The rationale for using such model was that the pathological trauma build up over time, reach a high point, then ultimately decline. Here, an ordinary least square was estimated $T = 0$ and two distributed lagged models were computed using ; $T = 10$ as in Gravelle (1981) and $T = 20$ as in the original Brenner (1979). Prior to the estimation, the series were transformed in log.

OLS (in log levels) :

	Coefficient	Std Error	T-Stat	P-value
Constant	8.959	0.22	40.55	0.001
UE	0.251	0.03	8.07	0.001
$R^2 = 0.99$ $DW = 0.59$ $F_{1,84} = 65.28\ddagger$				

\ddagger : significant at the 1% level.

The low value of the DW statistic suggested running the same regression in first differences instead of levels.

OLS (in first differences of log, i.e. in growth rates) :

	Coefficient	Std Error	T-Stat	P-value
Constant	41E-04	58E-03	0.07	0.001
dUE	0.232	0.059	3.93	0.001
$R^2 = 0.15$ $DW = 2.08$ $F_{1,83} = 17E - 04\ddagger$				

\ddagger : significant at the 1% level.

The polynomial distributed lag :

	Const.	$\hat{\lambda}_0$	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\sum_{i=0}^T \hat{\beta}_i$	T-Stat	R^2	F
$T = 10$	8.7424 \ddagger	0.0311 \ddagger	-0.0086 \ddagger	0.0011 \ddagger	0.2837 \ddagger	7.1931	0.999	$F_{(3,72)} = 19.51\ddagger$
$T = 20$	7.7378 \ddagger	0.0585 \ddagger	-0.0107 \ddagger	0.0005 \ddagger	0.4269 \ddagger	8.3717	0.999	$F_{(3,62)} = 25.71\ddagger$

\ddagger : significant at the 2% level.

A recursive residual estimates for - both distributed lag models - was estimated and concluded parameters stability over the period over 76:01 to 97:02.

3.5. Specification Using Bayesian Vector Autoregression (BVAR) :

There is no reason to suggest a restriction on the coefficients of a dynamic structure such as the one used in the polynomial distributed lag models. No "hard shape" restriction on the coefficients seems theoretical sound. Dealing with such problem, shrinkage estimators (on the lag coefficients) have been suggested and are more suitable for the analysis here. Dropping a lag from the dynamic structure is equivalent to forcing its coefficient to zero. Rather than the adopting a lag or no lag approach, one can suggest that coefficients on longer lags are likely to be close to zero than shorter lags. Meanwhile, if the data show evidence of a significant lag at longer time, one must allow for such event. In brief, one allow the data to decide by using a Bayesian framework in the dynamic structure. Litterman (1981) suggested a Bayesian framework to the VAR specification. Within such approach, one attempts to filter as much information form the data prior to the model specification, and let the data decide on the lag specification in the system. Using a symmetrical "atheoretical" prior, to decide which variable at which lag should be included, on all variables will balance the trade off between the overparametrization and oversimplification of the model. (*The prior acts as an antenna, when appropriately directed, may clarify the signal.*) The prior is "objective" in the sense that it is based on experience, that may reflect ignorance and have no economic interpretation. Write the VAR model as

$$y_t = Const + \Phi_1 \cdot y_{t-1} + \Phi_2 \cdot y_{t-2} + \dots + \Phi_p \cdot y_{t-p} + \varepsilon_t \quad (3.4)$$

where

$$\Phi_j = \begin{pmatrix} \phi_{11}^{(j)} & \phi_{1k}^{(j)} \\ \phi_{k1}^{(j)} & \phi_{kk}^{(j)} \end{pmatrix} \quad j = 1, \dots, p \quad (3.5)$$

The coefficient $\phi_{ik}^{(j)}$ gives the relation between y_{it} and $y_{k,t-j}$. ($i = 1, \dots, k$. $j = 1, \dots, p$) Note that y_t is a k -dimensional vector $y_t = (y_{1t}, \dots, y_{kt})'$ and ε_t is also a k -dimensional vector $\varepsilon_t = (\varepsilon_{1t}, \dots, \varepsilon_{kt})'$. In our context $y_t = (m_t, UE_t)'$ and $p = 8$. Different lags were estimated, $p = 4, 8, 12$ to capture the effect over 1, 2 and 3 years. The lag $p = 8$ was found most significant.

The Bayesian procedure is implemented by placing a Normal prior with mean zero on the coefficients of the lags⁸, and allowing smaller standard deviation the longer the lag is in time. Usually, a mean of one is placed on the first own lag,

⁸Except own first lag.

and means of zero on all other coefficients. (It center the prior around a random walk process) Formally, $\phi_{ii}^{(1)} = 1$ and all other $\phi_{ik}^{(j)} = 0$ to characterize the mean of the prior distribution of the coefficients.

Litterman assumed a diagonal variance-covariance matrix for the prior distribution, with γ referring to the standard deviation of the prior distribution for $\phi_{ii}^{(1)}$;

$$\phi_{ii}^{(1)} \sim N(1, \gamma^2) \quad (3.6)$$

For the other coefficients, the standard deviation of the prior decay with respect to the lag, i.e. more confidence that the value is zero for further lags.

$$\phi_{ik}^{(j)} \sim N(0, S^2(i, k, l)) \quad \text{for } i \neq k \quad (3.7)$$

Let k refer to the variables in the system ($k = m, ue$) and i refer to the equation whose dependent variable is i . Define the standard deviation of the prior distribution for lag l of the variable k in equation i as :

$$S(i, k, l) = \{\gamma \cdot g(l) \cdot f(i, k)\} \frac{s_i}{s_k} \quad (3.8)$$

where $f(i, i) = g(1) = 1.0$ such that $\phi_{ii}^{(1)} \sim N(1, \gamma^2)$ as above.

γ is the degree of overall tightness and represents the confidence in the prior information. A value of $\gamma = 0.2$ means that one have a confidence of 95% that $\phi_{ii}^{(1)}$ is no smaller than 0.6 and no greater than 1.4 (the mean is equal to 1). s_i is the standard deviation of the residuals of a univariate autoregression on the dependent variable of equation i (OLS of y_{it} on a constant and own p lags). s_i/s_k represent a correction for different scales of the variables. In other words, an adjustment for the units in which the data are measured. $g(l)$ is the tightness on lag l relative to lag 1. It capture how the standard deviation changes with increasing lags. $f(i, k)$ is the tightness on variable k in equation i relative to variable i , while γ represent the overall tightness. Standard function forms for $g(l)$ and $f(i, k)$ are :

$$f(i, k) = \left\{ \begin{array}{ll} \textit{symmetric} & f(i, k) = \left\{ \begin{array}{ll} 1 & i = k \\ w & i \neq k \end{array} \right\} \\ \textit{circular} & \textit{circle - star} \quad \textit{prior} \\ \textit{general} & \textit{in the case of 6+ variables} \end{array} \right\} \quad (3.9)$$

where w is a weight parameter, and represents the relative tightness applied to all off-diagonal variables in the system. To have more confidence in the prior belief that $\phi_{ik}^{(j)} = 0$ than the prior belief that $\phi_{ii}^{(1)} = 0$, w should be less than one. (a common choice in applied economic time series is $w = 0.5$ and $\gamma = 0.2$) As w goes to zero, the system reduces to as set of univariate autoregressions. In other words, forcing coefficients on other than own lags toward zero.

$$g(l) = \left\{ \begin{array}{ll} \textit{harmonic} & g(l) = \\ \textit{geometric} & g(l) = l^{-d} \end{array} \right\} \quad (3.10)$$

where d is the lag decay parameter. For the harmonic (geometric) function, a large (small) value for d reflect a tighter prior. Note that lag decay function $g(l)$ is a bad choice when present with seasonal data. So for the BVAR estimation, we smoothed the series by applying an additive seasonal filter on both, and exponential trend on the 'Death' series⁹. Criteria for selecting the method of smoothing were 'Sum of Squares' and the 'Schwarz Information'. In brief, a low tightness forces to VAR, while a high tightness forces to an OLS.

Many criteria for choosing the parameters of the prior are discussed in the literature. Among others, one can use the log determinant of the covariance matrix of out-of-sample forecasts errors or by computing a forecast performance statistic such as the Theil U statistic (the ratio of the root mean square error to the root mean square error of the naive forecast of no change in the dependent variable, i.e. forecasting no change over time). The later is used here.

The Bayesian approach is very flexible. There is neither restriction on lags, nor specification testing, and it allows different lags at different equations. The presence of trending variables does not cause any particular problems in this framework. Inference is based on the likelihood principle. The approach requires normality of residuals and good priors, but invariant to the size of the dominant root of the system. Estimation is carried out numerically, passing through the sample recursively with the Kalman filter algorithm.

When Σ and the coefficients on lags are not time varying parameters, the system of equations (3.6) and (3.7) forms a VAR model with a set of uncertain linear restrictions on the linear coefficients. The unconditional distribution of the coefficients are constants. If (3.7) is regarded as a dummy observation appended

⁹Different models were estimated and only these were selected among all combination of : 1) Trend; none, linear, exponential and 2) Seasonal; none, additive, multiplicative. Results are available.

to the system, then estimation of the model can be carried out with mixed-type estimation. The result is a restricted estimator which shrinks the data toward the information contained in the prior restriction. This interpretation of the BVAR is similar to the single equation ridge regressions: when the noise in the data is influential, the set of uncertain linear restrictions acts as a constraint on the filter extracting information from the data.

One might ask if this method will introduce an alternative source of bias such as shrinking the model to an incorrect parameter vector. A solution to this bias is to specify a prior distribution over different trend specifications, by trying different distributional assumptions on the innovation of the BVAR model, and computing pairwise posterior odd ratios using numerical integration. This approach may be useful for testing the model. We use the Theil U statistic to compare between models. We estimated different models¹⁰ using a range of priors, such that all combination between $\gamma = 0.1, 0.2, \dots, 2.0$ and $w = 0.1, 0.2, \dots, 1.0$. This statistic suggested that the model using $\gamma = 0.2$ and $w = 0.8$ reported the best forecast in the 'Death' series.

Forecast statistic for the series 'Death'

Step	Theil U
1	0.522
2	0.421
3	0.333
4	0.310
5	0.313
6	0.275
7	0.201
8	0.230

The Theil U statistic is a unit free measure, and provide a comparison with the naive (no change over time) forecast. A value higher than one, means that the model is doing worse than the naive one.

3.6. Impulse responses and variance decomposition :

Both are methods to describe the dynamic properties of the model following certain shocks. In our case, the shock is a recession (i.e. a high UE shock to the

¹⁰Results are available and not reported for space use.

economy). Both, are in-sample forecasting exercises. They are isomorphic in respect to the information they contain. The former is used to better economic understanding, while the later is used to economic testing. They describe the effect on the system of equations of a "typical" shock to a variable, where "typical" is used in the sense of a one standard error shock. So a shock to UE - a recession - effects on mortality can be traced as a mean to quantify the Brenner hypothesis. In general :

Suppose y_t is a covariance stationary process (possibly after some transformation) with MA representation.

$$\begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = \begin{pmatrix} D_{11}(l) & D_{12}(l) \\ D_{21}(l) & D_{22}(l) \end{pmatrix} \begin{pmatrix} w_{1t} \\ w_{2t} \end{pmatrix} + \begin{pmatrix} B_{11}(l) & B_{12}(l) \\ B_{21}(l) & B_{22}(l) \end{pmatrix} \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix} \quad (3.11)$$

For such an y_t , Σ is in general non-diagonal.. To transform the system in such a way that shocks to each equations are uncorrelated one can use a Cholesky decomposition. Given that Σ is real symmetric positive definite matrix, then let L be a non-singular lower triangular orthogonal matrix with ones on the main diagonal and let V be a diagonal matrix. If we choose $L^{-1}.V.(L^{-1})' = \Sigma$, then the system can be normalized as :

$$\begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = \begin{pmatrix} D_{11}(l) & D_{12}(l) \\ D_{21}(l) & D_{22}(l) \end{pmatrix} \begin{pmatrix} w_{1t} \\ w_{2t} \end{pmatrix} + \begin{pmatrix} G_{11}(l) & G_{12}(l) \\ G_{21}(l) & G_{22}(l) \end{pmatrix} \begin{pmatrix} v_{1t} \\ v_{2t} \end{pmatrix} \quad (3.12)$$

Where $G(l) = B(l).L$ and $v_t = L^{-1}\varepsilon_t$. Because L is lower triangular, so is $G(l)$ and innovations in variable i do not contemporaneously affect variable k if variable k precedes variable i in the list of elements of y_t . Note that $L^{-1}.V^{1/2}$ has the standard deviation of ε_t along its principal diagonal. Thus a shock of one unit to v_t is equivalent to a shock of one standard deviation to ε_t . The VAR system (3.12) has a so-called Wold causal chain form. Note, that this orthogonalization procedure is not unique and depend on the ordering of the variables, i.e. the position of each variable in the y_t vector. (simply interchange the rows or columns in Σ , you will get a different Cholesky factor)

The variables on the top of the triangle contemporaneously feed into all the other variables and the variables on the bottom of the triangle contemporaneously affect only themselves.

1) Variance decomposition tells us how much of the average squared forecast error variance of one variable at the k th step ahead is associated with surprise movements in each variable of the model.

2) The impulse response function traces out the moving average representation of the system and describes how one variable responds over time to a single surprise increase in itself or in any other variable.

If the errors of the system are uncorrelated, then an impulse response function (IRP) interpretation is straightforward. It describes the response of mortality to a shock in the innovations. If the errors of the system are correlated, then the IRP interpretation is ambiguous. The errors have a common component which can not be identified with any specific variable. The IRP will reflect the response to the innovation shock and to the common component. For the model to be meaningful, one need to transform the errors of the system into a contemporaneously uncorrelated form. Why ? Simply to be able to distinguish the response of mortality to an UE shock. To transform the VAR innovations into an orthogonal form, one use a Cholesky decomposition of Σ (the var-cov matrix of the errors).

IMPULSE RESPONSE :

The matrix $G(l)$ in (3.12) represents the impulse response functions and is useful in examining the effects of typical shocks to the variables of the system in the short and long run.

Each G_j describes the response of the vector y_t to innovations j periods ago. The k th row of each G_j measures the responses of y_{kt} to innovations in the system which occurred j periods ago, $j = 0, 1, 2, \dots$. Finally, the h th element of the k th row of $G(1)$ measures the cumulative effect on y_{kt} of an innovation in y_{ht} that occurred j periods ago, where $j \rightarrow \infty$.

VARIANCE DECOMPOSITION :

To compute the variance decomposition of y_t note that from (3.12)

$$var(y_t) = G(1)Ev_t v_t' G(1)' \tag{3.13}$$

Special case; $m_1 = m_2 = 1$. Then, the variance of y_{1t} has two components: one due to the impact of its own innovations from time t to time $t-j$, $j = 1, 2, \dots$: and one due to innovations in y_{2t} from time t to time $t-j$. If y_{2t} is shocked at time $t-j$ and left unperturbed afterward, we can examine how much of the variability of y_t at time t is due to that innovation, for all j . As seen from (3.13) the variance decomposition adds no new information to the impulse response analysis, but instead, presents it in an alternative form.

The impulse response function to a shock of one standard deviation in the series 'Unemployment' using a VAR classical estimation and BVAR - where $\gamma = 0.2$ and $w = 0.8$ - respectively..

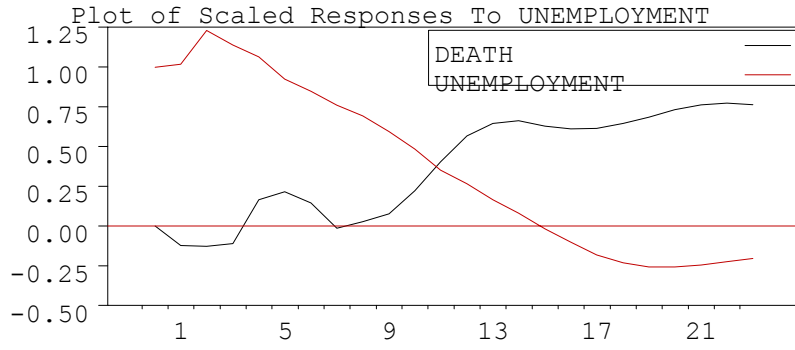


Figure 3.2:

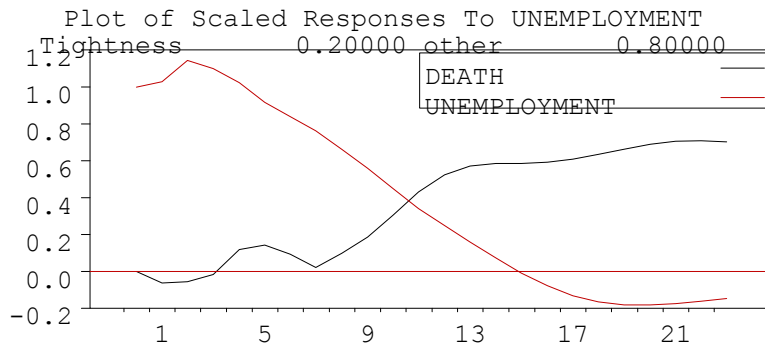


Figure 3.3:

And the variance decomposition of the forecast in 'Death' follows.

Step	UE	Death	Step	UE	Death
1	0.25	99.75	8	1.70	98.30
2	0.14	99.86	9	1.97	98.03
3	0.11	99.89	10	2.87	97.13
4	0.10	99.90	11	4.94	95.06
5	0.71	99.29	12	8.32	91.68
6	1.43	98.57	13	12.47	87.53
7	1.72	98.28	14	16.64	83.36

4. Extensions and Conclusion :

One can attempt to use micro-data and test the probability of low skilled workers being unemployed effect on their mortality. Such correlation - if found - would definitely reinforce the Brenner hypothesis. Also, micro-data on causes of death and other morbidity data correlation to unemployment, such as the relation between first hospital visit, the increase of incidence of acute mental disorder might reveal a missing link in the chain of effects between high unemployment and mortality. Such agenda will reveal what type of reason¹¹ is responsible for the hypothesis. Also, one may attempt to test if the relationship is stable across countries ?

This study concludes that following a recession shock (an increase of unemployment), mortality will reach a peak after 6 quarters and will trend up after 8 quarters. Exactly, by how much will that be. We estimate that 8.32% of the variance in mortality is due to the shock in unemployment after 12 quarters increasing to 16.64% after 14 quarters.

5. References :

References

- [1] Brenner, M. Harvey. (1979) "Mortality and the national economy : a review and the experience of England and Wales 1936-76" *Lancet*, 15 September, pp 568-73.

¹¹Three reasons were discussed in section I.

- [2] Gravelle H.S.E., Hutchinson G. and Stern J. (1981) "Mortality and unemployment: a critique of Brenner's time series analysis" *Lancet*, 26 September, pp 675-9.
- [3] Hamilton, James D. (1994) "Time series analysis" *Princeton University Press*.
- [4] Kasl, Stanislav V. (1979) "Mortality and the business cycle : Some questions about research strategies when utilizing macro-social and ecological data" *American Journal of Public Health*, Vol 69, No 8, pp 784-8.
- [5] Litterman, R (1981) "A Bayesian procedure for forecasting with vector autoregressions" *Federal Reserve Bank of Minneapolis*, working paper.