RESERVE
BANK
O F   N E W   Z E A L A N D

DP2000/07

# Stylised facts from output gap measures

## Alasdair Scott

## January 2000

## JEL classification: E32

## Discussion Paper Series

## Abstract

This paper compares three models of the output gap in New Zealand – the Reserve Bank of New Zealand's incumbent MV filter, estimates from a Structural VAR, and a multivariate unobserved components model – and investigates whether there are features that are consistent across the measures of the gap. Various detrending methods are used for benchmarking the output gap measures, including a linear trend, a fourth difference filter, a band-pass filter, the Hodrick-Prescott filter, an "optimal $\lambda$" procedure, and a nonparametric estimator of permanent trend. The estimates of the gap are examined to see how they compare as to the lengths and amplitudes of cycles, whether they exhibit regular periodicity and regular shapes, whether they are symmetric in phases and severity of swings, and whether they point to the same turning points. The analysis leads to the conclusion that while different filters result in estimates of the output gap with quite different properties, the three models are generally in agreement about the historical profile of the output gap, if not its precise level. Moreover, there are signs that the models are increasingly in agreement about the level of the gap, indicating that the growth cycle is becoming more regular in the 1990s.

# 1    Introduction

It is common to approximate the level of potential output with an estimate of the underlying trend.  However, detrending is non-trivial – different methods can be expected to reveal different aspects of the data.[1]  This paper asks whether there are stylised facts from different models of the output gap in New Zealand.  The exercise is constructed as a "shooting match", where different estimates of the output gap over the same data set are compared using a variety of simple statistical criteria and tests.  I am interested in what, if anything, is said in common from different estimates of the output gaps.[2]

I use various means of comparison.  The statistics are deliberately simple and non-parametric.[3]  Peaks and troughs in the detrended series are identified using a simple algorithm, and this allows us to measure the phases of expansion and contraction. Simple statistics for average durations, average amplitudes, and average amplitudes per period are calculated, followed by a test for duration dependence.  I follow with stylised facts of the symmetry of periods and amplitudes.  Finally, I ask whether the measures co-move, using correlation analysis and concordance statistics.  These latter statistics are very relevant to policy, since they focus on the turning points from positive to negative gaps that are arguably most essential for a policymaker.

Three models of the output gap are compared against benchmark detrending methods. These detrending methods are well-understood, with known properties, so that we can hope that some insight into the properties of the more complicated models can be gleaned by comparison to the simple models.  The selection of detrending methods is therefore deliberately broad, and includes non-parametric and parametric methods, local and global models, and both one-sided and two-sided filters.  One unfortunate problem faced is that the more complicated methods are applicable only over a recent sample, which means that inference about the properties of the output gap has to be done with care.

Nonetheless, the following points stand out from the exercise.  First, the estimates of the output gap over New Zealand output data represent an approximate average of the benchmark methods in terms of the number of cycles identified over the sample.  They are also approximately symmetric – they identify roughly half of the sample period as a state of excess demand and half as excess supply.  This illustrates the essentially "two-sided" nature of the trends that are estimated from the models.  There is no evidence of a regular duration to expansions and contractions, but there is evidence of a typical "shape" to the cycles – statistically, the rates of expansion and contraction both tend to be regular.  There is large disagreement about the swing of cycles, however – the typical cumulative decline, in particular, varies widely.  Nonetheless,

---

[1]     For an extensive discussion, see Canova (1998).

[2]     Similar exercises have recently appeared for U.S. data: see, inter alia, Canova (1999).

[3]     Stylised facts from non-parametric tests are useful for informing the desired quality of a parametric time-series model, whereas to move straight to parametric modelling of the gaps would be to beg the question of their properties.

all the measures are highly correlated, and typically agree on the turning points of a transition from phase to phase (that is, negative and positive output gaps) about 80 percent of the time.

The outline of the paper is straightforward. In section 2, the different detrending methods and the models of the output gaps are described briefly. In section 3, the tests are explained and the results for all the gap measures are presented. Section 4 includes some remarks on the significance of the results for policy and future research.

## 2      Different measures of the gap

The detrending methods used in this exercise are deliberately quite different. The methods range in complexity, and include non-parametric and parametric methods, local and global models, and both one-sided and two-sided filters. In addition to discussing their estimation, in the following I discuss the implicit identification schemes implied by the detrending methods. As will become apparent, different methods contain quite different priors about the evolution of output.

Before proceeding, a note on notation and language is in order. In the discussion that follows, $y_t$ will refer to an observation at date $t$ from the output series. Trends are labelled $\mu$, and it is assumed that we are able to use the terms "trend" and "potential output" interchangeably.

I examine three models of the output gap: the Reserve Bank's MV filter (Conway and Hunt, 1997), the results from a Structural VAR (Claus, 2000), and a multivariate unobserved components approach (Scott, 2000). These are compared against the following benchmarking methods: linear and segmented trends, a fourth difference filter, the band pass filter of Baxter and King (1995), the Hodrick-Prescott (1997) filter, an "optimal $\lambda$" filter (Agénor, McDermott and Prasad (2000)), and the nonparametric estimate of trend used in Coe and McDermott (1997).

A linear trend (LIN) is one of the easiest and earliest methods used to estimate potential. Here $\mu_t$ is determined by the regression of $y_t$ on $t$. The method implicitly assumes that potential output is deterministic with a constant growth rate. This identification implies that all shocks are demand shocks, thereby matching a traditional Keynesian assumption that in "mature economies" contractions in output are problems of insufficient demand.[4] Although still in current use, the method therefore ignores more recent theoretical arguments that $\mu_t$ be treated as a stochastic process. Nor does it match the observed tendency for output growth rates to decline, raising the possibility of quite unrealistically optimistic estimates of potential as time goes by.

The use of a segmented trend (SEG) is a somewhat ad hoc attempt to solve constant growth rate assumption problem of the linear trend. It remains an easy method to implement, but requires a known break point. The identification of breakpoints is

---

[4]      See Tobin, (1997).

nontrivial: they are often contentious and are statistically very hard to identify.[5] Further, it may take time for the effects of even well-known discrete events to diffuse through an economy. For the purposes of illustration, I impose a break for New Zealand output at 1985:2, which corresponds roughly to an exchange rate regime shift and the beginning of an extensive reform process.

Difference filters assume that the underlying trend follows a random walk; the trend is therefore simply the previous observation, $y_{t-p}$, where $p$ is the order of the filter. As compared to the previous two methods, the trend is now stochastic, but in the presence of on-average growth, most shocks are implicitly identified as demand shocks. This may be a reasonable characterisation of policy, at least traditionally.[6] However, such a filter is a poor match to conventional preconceptions of business cycle frequencies, since the transfer function amplifies high frequencies while attenuating low and business cycle frequencies (see figure 1). On the other hand, annual growth rates are often used as a proxy for pressure on productive capacity, and so the fourth difference filter (DIFF4) is used here. This might be seen as an attempt to improve on the high frequency "noisy transfer" problem of the first difference filter. By labelling the trend as the observation four quarters previously, the filter captures more of the low frequency movements in the data. A further perceived advantage of the growth rate is that it is a one-sided filter, and therefore not subject to revision. However, the transfer function is still far from ideal, as seen in figure 1: the filter is nowhere an ideal "unity gain" transfer, and in particular there is zero gain at the annual frequency.[7]

The band-pass filter (BPF) of Baxter and King (2000) deals directly with the frequency-domain properties of the data. The aim is to eliminate both high-frequency irregular components of the data and zero- to low-frequency components, thereby isolating business cycle frequencies. The filter passes through business cycles frequencies (usually in the range of 6 to 32 quarters as per the NBER business cycle classification) which are then identified with demand-led movements. Rather than use spectral masking techniques that can result in cross-spectral irregularities, Baxter and King implement the filter in the time domain with a $k^{th}$-order two-sided moving average process. This results in a filter that for the infinite-sample case has ideal unity gain on business cycle frequencies and zero gain on all others. However, in practice, when $k < \infty$ the bands will "leak". (Figure 1 shows a comparison of the transfer function of the ideal filter and the filter with $k = 12$.)

The trend from a Hodrick-Prescott (1997) filter (HP) can be thought of as the solution to an optimisation problem involving minimising both deviations of the estimated trend from the observed data and squared movements in the estimated trend. The balance between these two objectives is determined by the parameter $\lambda$. As $\lambda$ tends to infinity, the result is a linear trend; as $\lambda \to 0$ the estimated trend will follow the data more and more exactly. The filter therefore encapsulates different theoretical priors on the equilibrium level of output. Correspondingly, the ratio $\lambda$ can be thought of as a prior on the ratio of demand to supply shocks. (Here $\lambda$ takes the conventional

---

[5]    See, for example, Perron (1989).

[6]    For example, Okun's law was originally specified as a relation between changes to unemployment and the growth rate of output. See Okun (1962).

[7]    In radians, 1.57.

value of 1600.) The disadvantage of using this filter is therefore primarily that the answer to one of the main questions of interest is "dialled in" by construction. On statistical grounds, the assumptions of the HP filter may not be justified by a given data set.[8] The HP filter is also known to induce spurious cycles from a random walk.[9] On the other hand, the filter is easy to use, reasonably robust, and is certainly commonly used.

Noting the problems of the HP filter, Agénor, McDermott and Prasad (2000) propose a method for estimating the degree of smoothing, which I will refer to as the "optimal $\lambda$" filter (OPTL). The procedure employs a data-dependent selection criterion which can be thought of as minimising an in-sample forecasting error when the value of $y_t$ is (deliberately) unknown. This procedure will produce a smoothly-evolving trend that has an exact equivalent in terms of a given $\lambda$ for the HP filter.

The final benchmark method is a nonparametric estimate of the underlying trend employed by Coe and McDermott (1997). This technique uses a univariate nonparametric regression method to estimate the trend and cyclical components of a series without having to specify the functional form of the trend component of the underlying series or the degree of smoothing applied to the actual data. The advantage of these last two methods is that they select optimal trends (in the sense of satisfying a statistical fit criterion) without the need for assumptions or tweaks by the user.

This is not an exhaustive list. Other univariate methods that have been proposed include the Beveridge-Nelson (1981) decomposition, the unobserved components models of Watson (1986) and Harvey and Jaeger (1993), and the Markov-switching framework of Hamilton (1989). The methods used here, however, are all mechanical and known to be robust.[10]

The following multivariate methods include structural relationships to identify the gap. They too are variously non-parametric and parametric.

The MV Filter (MV) is the Reserve Bank of New Zealand's incumbent measure of potential.[11] Taking the HP filter as its core smoother, it can be viewed as an attempt to introduce economic structure into the measure of the gap by the addition of conditioning relationships such as an Okun's relation, a Phillips curve, and a capacity equation.[12] These relationships and their parameters are not estimated as a system. In practice, the filter is considered to make more sense than HP-filtered estimates while

---

[8]     See Harvey and Jaeger (1993).

[9]     See Cogley and Nason (1995).

[10]    Experimentation with the Harvey and Jaeger (1993) trend plus cycle and the Beveridge-Nelson (1981) decomposition using an ARMA model was not successful.

[11]    See Conway and Hunt (1997).

[12]    The current MV filter used at the Reserve Bank of New Zealand uses a survey measure of inflation expectations in the Phillips curve. For the purposes of this exercise, the estimate of the gap has been back-dated to 1971:1 by using a one-sided moving average of inflation as a proxy for expectations up to 1982:4.

remaining robust in real time, an essential feature for an estimate of the trend to be used for forecasts.

A systems estimator of the output gap in a multivariate setting is provided by the Structural VAR (SVAR) of Claus (2000). This estimate uses the standard Blanchard and Quah (1989) method to identify the gap from a system consisting of output, employment, and capacity utilisation data. The gap is identified on the basis of a structural decomposition of innovations into permanent and transitory components.

Another systems estimator comes from the multivariate unobserved components approach (UC) of Scott (2000). This model is an extension of the models of Watson (1986) and Clark (1989) to introduce a production-function approach in a structural time series model setting with relationships of unemployment (labour) and capacity utilisation (capital) to common cycle in output. The relations are estimated simultaneously along with the parameters for the stochastic trends by the Kalman filter and exact maximum likelihood.[13] While the data are left relatively free to speak about the properties of the trends, the common cycle, and the Okun's law and capacity equation, the method is quite demanding on the data.

The data used are real, seasonally-adjusted production GDP for all methods. The Household Labour Force Survey unemployment rate is used in the MV and UC approaches, while Household Labour Force survey employment is used for the SVAR estimates. An index of capacity utilisation is utilised in the MV and UC models. The inflation rate used in the MV filter is the annual rate of change in the CPI ex interest index.

## 3    Results

The comparison involves assessing the estimates by a series of nonparametric statistics. Essentially, we are interested in answering a series of questions to see whether there are any robust stylised facts that come through, in spite of the differences between the methods and their implications. First, what does a typical gap cycle look like in terms of duration and amplitude? Second, are the gap cycles periodic – does the probability of moving to an expansion (contraction) increase as the contraction continues? Third, are the gaps asymmetric, both in terms of the duration of contractions and expansions, and in terms of the severity of contractions and expansions. Finally, do the measures co-move, and hence provide the same signals to the policymaker?

In order to facilitate this comparison, a method is needed to date the phases of the gap estimates. Here, a simple algorithm is applied which dates a peak as the highest point in the period during which output is above its trend (that is, the gap is positive). The opposite applies to date troughs. This is a very naïve dating rule, and it could be convincingly argued that the policymaker would never follow such a rule in practice. For example, this rule dates a peak of an excess demand phase no matter how close that peak is to zero; in practice, it seems likely that a policymaker would attempt to

---

[13]    The final estimate of the gap from the UC method is obtained from the Kalman smoother. See Harvey (1990).

distinguish between a phase that was "close enough" to neutral and a phase that was "clearly" in excess demand. Implementing such a decision rule is non-trivial. Two alternative rules were experimented with. In the first rule, a three quarter one-sided moving average filter was applied to the gap estimates, simulating the actions of a policymaker who is attempting to filter out high-frequency spuriae from their decision-making problem. This did not change the results reported below substantially. A second rule eliminated peaks and troughs that were within a certain distance to zero. However, quite aside from the subjectivity of the threshold rule required, further ad hoc filtering is required using this rule in order to ensure that peaks and troughs alternate. Experience with the gap estimates revealed that it was very difficult to know when to stop this censoring process. The results would then be susceptible to the critique that the stylised facts had, in essence, been "dialled in". For these reasons it was decided to proceed with the simple dating rule.

The estimates of the gaps and the datings of the peaks and troughs can be seen in figures 2 and 3. As can be quite clearly seen, the various methods imply very different estimates of the gap. We will quantify those impressions in the analysis below.

**Average durations and amplitudes:**

A first pass at the stylised facts of the gaps is performed by calculating the average durations, amplitudes, and amplitudes per quarter of the phase for both contractions and expansions. (For this purpose, contractions are defined as the phase from peak to trough, and expansions as the phase from trough to peak.) These are presented in table 1a. For each of the measures, there is a wide variety of averages between the gap estimates: for example, in the first column of table 1a the durations range from -4.73 quarters for the OPTL technique to -12.67 for linear detrending. However, the output gap models cluster around the upper end of the range at 8 to 12 quarters for contractions and 8-11 quarters for expansions.[14]

There is also a wide range for amplitudes. To a large extent, this should be expected, given the different properties of the detrending methods. The output gap models are in quite tight agreement of about 5 percentage points for both expansions and contractions. Overall, the estimates appear to be placed between the two extremes of a straight line (LIN) and following the data closely (BPF).

**Symmetry and severity:**

All of the estimates are approximately symmetric, with one exception. For each measure, durations, amplitudes and quarterly amplitudes are approximately equal for contractions and expansions. This is largely to be expected from the methods, since – with the exception of the fourth difference filter – involve two-sided moving averages to determine the trend. A complimentary measure is provided by the second column of table 1b, which measures the percentage of time that the series spend in an expansion phase. There the asymmetric nature of DIFF4 is seen clearly: 77 percent of time, output is spent in a state of excess demand. This is *not* a desirable feature from a model of the output gap, as it would put the monetary policymaker in a tightening bias

---

[14] One can also see that the numbers for the MV filter are pulled down by a number of peaks and troughs in the middle of the sample that could be judged to be spurious.

most of the time. On the other hand, the output gap measures are symmetric; the UC model achieves the desirable figure of 50 percent.

The price to be paid for this symmetry, however, is that all of these estimates (apart from DIFF4) will be subject to revision. To the extent that the filters achieve a more accurate picture of the underlying trend by using information over the whole sample, the estimates at $t$ will be subject to revision as more and more data are added to the sample. The extreme case occurs in the band-pass filter, for which the estimate of the trend is not defined for the last $k$ (here 12) observations.

Two measures of severity are provided. First, columns three and six of table 1b state, respectively, the maximum amplitude of the contraction phases and expansion phases identified by each method. The estimates vary tremendously. The fourth and seventh columns of the table present measures of, respectively, the average cumulative loss (gain) during contractions (expansions) for each measure. These numbers also indicate the difference between the one- and two-sided views of the world: for DIFF4 the average cumulative loss of contractions is 7 percent whereas the average cumulative gain of expansions is 30 percent.

For the other measures, their cumulative gains and losses tend to be similar. However, there is wide disparity between measures about the severity of contractions and expansions – the figure for LIN is nearly ten times the figure for BPF. The output gap measures, the figure for SVAR is nearly twice that of the benchmark HP1600. This implies that the levels of the gaps are estimated quite imprecisely, suggesting that the measures should be used as indicators primarily of "which way the wind is blowing" rather than as certain estimates of the level of the gap.

**Duration dependence and rank correlation**

Given the symmetry figures, a natural question that arises is whether the phases are somehow periodic. This can be tested by the application of the Brain-Shapiro (1983) test for duration dependence, which expresses the idea that the longer the series remains in expansion (contraction), the more likely it is to switch to a contractionary (expansionary) phase. The null hypothesis is that this probability is independent of the length of the phase and the statistics are presented in the fourth and eighth columns of table 1a. The picture is uniform across measures and phases. There is no evidence of duration dependence from any of the measures, and hence no evidence of a regular periodic cycle.[15] Given the changing economic environment over the sample period, this is probably to be expected.

Columns 5 and 8 of table 1b present Spearman rank correlation statistics. Applied to contractionary phases ($corr_{P-T}$), this tests the hypothesis that the depth of the contraction is significantly correlated with the duration of the contraction. The test can therefore be interpreted as a test for a regular "shape" in terms of the triangular properties of the cycle. Whereas there is little evidence for a regular period of expansions and contractions from any of the measures, there is some evidence for a

---

[15] The statistic for duration dependence is distributed $N(0,1)$; the 5 percent critical value for the two-sided test is therefore 1.96.

regular shape in contractions and wide evidence of a regular shape in expansions. It appears to be a particularly strong feature of the SVAR and UC models.[16]

**Correlation and concordance statistics:**

We are interested in the question of whether the measures of the gaps co-move with each other, and specifically interested in whether the measures give essentially the same signals about whether the economy is above or below potential. To analyse this, I make use of two statistics, the familiar correlation statistic and the less well-known concordance statistic.

The correlation matrix is presented in table 2. All of the 45 combinations are significant at the five percent level or greater.[17] On this basis, we might be led to be indifferent about which measure is used, since apparently they almost always "co-move". However, correlation measures essentially mix amplitudes and durations measures together; as shown by McDermott and Scott (1999), the covariance of two series may be dominated by the amplitude of particularly large swing that is common to both series. For the policymaker, it may be more relevant to know whether the measures signal whether the economy is above or below potential at the same time or not.

For this purpose, we make use of the concordance statistic originally proposed by Pagan and Harding (1999) and examined in McDermott and Scott. This is a simple non-parametric statistic that measures the proportion of time two series, $x_i$ and $x_j$, are in the same state. Let $\{S_{i,t}\}$ be a series taking the value unity when the gap measure $x_i$ is positive and zero when it is negative. Define the series $\{S_{j,t}\}$ in the same way. The degree of concordance is then

$$C_{ij} = T^{-1}\left\{\sum\left(S_{i,t} \cdot S_{j,t}\right) + \left(1 - S_{i,t}\right)\cdot\left(1 - S_{j,t}\right)\right\},$$

where $T$ is the sample size. As a proportion, the values that the expression (1) may take are clearly bounded between zero and one. Faced with an empirical result of, say, 0.7, it is natural to assume that this is a large number relative to zero. However, in this case there is an equal chance that the series are in or out of phase with each other, and the distribution of the statistic will be symmetric around 0.5.[18]

Results from the concordance analysis are presented in table 3. All but two of the combinations are significantly different from the null of a random result of 0.5. The

---

[16]   For contractions, the figures for SVAR and UC are both at their upper limit of 1. However, this is still only significant at the 5 percent level, due to the limited data span and hence number of cycles in the sample period.

[17]   The significance levels are given by $1.96\times 1/\sqrt{T}$ for the 5 percent level and $2.58\times 1/\sqrt{T}$ for the 1 percent level.

[18]   It follows that if the two series $x_1$ and $x_2$ were independent, then the variance of the concordance statistic would be $1/[4(T-1)]$, where $T$ is the sample size of $x_1$ and $x_2$. The critical values follow accordingly.

average for the table is 78 percent. Clearly, related methods result in gap measures that are highly concordant: MV is 80 percent concordant with OPTL and 79 percent concordant with HP1600, for example. In particular, the concordances between MV, SVAR, UC and the benchmark HP measures range from 67 to 87 percent. This indicates that the measures do tend to provide the same signalling information – while the measures may be imprecise about the level of the gap, they tend to indicate the same periods of excess demand and excess supply.

One of the benefits of using the concordance statistic as a measure of co-movement is that it can be depicted – that is, since it is a summary of the proportion of time that two series spend in the same phase, we can show this "in-phase" behaviour over time. Figure 4 plots the MV and UC gaps next to a "bar code" which depicts this behaviour. The bar code is solid when the two series are in the same phase and blank when they are out of phase. Similar graphics are repeated in figures 5 and 6 for MV and SVAR gaps and UC and SVAR gaps. It is noticeable that the measures tend to be in-phase over the 1970s and in the 1990s but not in the 1980s. This indicates that the period of macroeconomic reform is problematic for the estimation of output gaps. It also implies that the economy has returned to a more conventional business cycle pattern, which will make the identification of permanent and temporary components of output easier in the future.

# 4 Conclusion

This paper sets up a "shooting match", where three measures of the output gap are benchmarked against various simple univariate detrending methods. Following a simple dating rule to identify peaks and troughs of the growth cycles in the detrended data, a number of nonparametric statistics were used to assemble some stylised facts about the gap measures. It was shown that the output gap methods are symmetric, acting as two-sided filters. For the policymaker, this is a desirable feature. However, it also has the disadvantage that estimates of the gap will be revised as more and more data become available. In this respect, these more sophisticated models are no panacea to the problems of two-sided filters. The measures will also tend to impose the symmetry property onto the measure of the business cycle, whereas classical cycles tend to be asymmetric.[19]

The three measures agree approximately about the average durations of cycles and tend to co-move. Like the benchmark detrending methods, however, they disagree about the severity of swings in the cycles. This uncertainty implies that the gap is best treated as an indicator of the state of the world rather than an exact measure of the precise level of the output gap. However, there is some evidence that this is, in large part, due to the influence of the sample period over the 1980s, where the estimation of the output gap seems particularly problematic. Over the 1990s the estimates have tended to converge, which may imply the emergence of a more regular growth cycle. All else being equal, this will be more conducive to more precise estimates of the underlying trend and the temporary component of output as time goes by.

---

[19]    See McDermott and Scott (1999).

# References

Agénor, Richard, C John McDermott and Eswar Prasad (2000), "Macroeconomic fluctuations in developing countries: some stylised facts." *World Bank Economic Review*, forthcoming.

Baxter, Marianne, and Robert King (1995), "Measuring business cycles: approximate band-pass filters for economic time series." *National Bureau of Economic Research Working Paper* No 5022.

Beveridge and Charles Nelson (1981), "A new approach to the decomposition of economic time series into permanent and transitory components with particular attention to the business cycle" *Journal of Monetary Economics* 7, 151-74.

Blanchard, Olivier and Danny Quah (1989), "The dynamic effects of aggregate supply and demand disturbances" *American Economic Review* 79, 655-73.

Brain, Carlos W and Samuel S Shapiro (1983), "A regression test for exponentiality: censored and incomplete samples." *Technometrics* 25, pp.69-76.

Canova, Fabio (1998), "Detrending and business cycle facts." *Journal of Monetary Economics* 41, 475-512.

Canova, Fabio (1999), "Does detrending matter for the determination of the reference cycle and the selection of turning points?" *The Economic Journal* 109 (January), 126-50.

Clark, Peter (1989), "Trend reversion in real output and unemployment" *Journal of Econometrics* 40, 15-32.

Claus, Iris (2000), "Estimating potential output: a structural VAR approach." manuscript, Reserve Bank of New Zealand.

Cogley, Timothy, and J Nason (1995), "The effects of the HP filter on integrated series." *Journal of Economic Dynamics and Control* 19, 253-278.

Coe, David T. and C John McDermott (1997), "Does the gap model work in Asia?" *International Monetary Fund Staff Papers* 44(1), 59-80.

Conway, Paul, and Benjamin Hunt (1997), "Estimating potential output: a semi-structural approach." *Reserve Bank of New Zealand Discussion Paper G97/7*.

Hamilton, James (1989), "A new approach to the econometric analysis of nonstationary time series and the business cycle." *Econometrica* 57, 357-84.

Harvey, Andrew C (1990), *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge University Press.

Harvey, Andrew C, and Albert Jaeger (1993), "Detrending, stylised facts and the business cycle." *Journal of Applied Econometrics* 8, 231-247.

Hodrick, Robert, and Edward Prescott (1997), "Post-war US business cycles: an empirical investigation." *Journal of Money, Credit and Banking* 29(1).

Kim, Kunhong, Robert A Buckle, and Vivian B Hall (1994), "Dating New Zealand business cycles" *New Zealand Economic Papers.*

McDermott, C John, and Alasdair Scott (1999), "Concordance in business cycles." Reserve Bank of New Zealand *Discussion Paper G99/7.*

Okun, Arthur (1962), "Potential GNP: its measurement and significance." Proceedings of the Business and Economics Section, American Statistical Association 1962, 145-58.

Pagan, Adrian, and Don Harding (1999), "Knowing the cycle." Manuscript, University of Melbourne.

Perron, Pierre (1989), "The Great Crash, the oil price shock, and the unit root hypothesis." *Econometrica* 57, 1361-1401.

Scott, Alasdair (2000), "A multivariate unobserved components model of the output gap." manuscript, Reserve Bank of New Zealand.

Tobin, James (1997), "Supply constraints on employment and output: NAIRU versus natural rate." *Cowles Foundation Discussion Paper* No 1150, Yale University.

Watson, Mark (1986), "Univariate detrending methods with stochastic trends" *Journal of Monetary Economics* 18, 49-75.

**Table 1a:**
**Descriptive statistics**

| | Peak to trough | | | | Trough to peak | | | |
|---|---|---|---|---|---|---|---|---|
| | Dur. | Amp. | Amp./ quarter | Brain- Shapiro | Dur. | Amp. | Amp./ quarter | Brain- Shapiro |
| LIN | 12.67 | -10.06 | -1.00 | 0.21 | 16.50 | 9.72 | 0.59 | -0.31 |
| SEG | 8.00 | -7.98 | -1.15 | 0.09 | 14.40 | 8.16 | 0.58 | -1.63 |
| DIFF4 | 7.88 | -7.18 | -1.13 | 0.86 | 5.25 | 6.68 | 1.71 | 1.57 |
| BPF | 7.00 | -1.57 | -0.34 | 0.06 | 5.75 | 1.64 | 0.47 | 0.95 |
| HP1600 | 7.00 | -4.67 | -1.13 | -0.63 | 7.50 | 4.45 | 0.80 | 0.89 |
| OPTL | 4.73 | -3.16 | -1.02 | 1.38 | 4.75 | 3.18 | 0.87 | 1.63 |
| NONPA RA | 8.00 | -7.45 | -1.10 | 0.09 | 14.40 | 8.53 | 0.62 | -1.63 |
| MV | 7.50 | -4.29 | -0.74 | -0.46 | 7.57 | 4.31 | 0.85 | 0.50 |
| SVAR | 9.00 | -5.32 | -0.64 | 0.76 | 10.67 | 4.86 | 0.47 | 0.00 |
| UC | 11.80 | -5.50 | -0.53 | 0.46 | 8.50 | 4.94 | 0.54 | 0.14 |
| Averages | 8.39 | -5.75 | -0.88 | 0.28 | 9.53 | 5.69 | 0.75 | 0.21 |

† denotes significance of the concordance statistic at the five percent level.
‡ denotes significance of the concordance statistic at the one percent level.

**Table 1b:**
**Descriptive statistics**

| | # cycles | % exp. | Peak to trough max. | loss | corr$_{P\text{-}T}$ | Trough to peak max. | gain | corr$_{T\text{-}P}$ |
|---|---|---|---|---|---|---|---|---|
| LIN | 3 | 50 | -13.26 | -46.84 | -0.50 | 13.13 | 51.18 | 0.40 |
| SEG | 4 | 54 | -12.82 | -35.56 | -0.40 | 12.33 | 40.05 | 0.90† |
| DIFF4 | 7 | 77 | -17.21 | -6.87 | -0.81† | 15.38 | 29.95 | 0.71† |
| BPF | 7 | 46 | -4.22 | -4.99 | -0.86† | 3.31 | 5.64 | 0.71† |
| HP1600 | 7 | 56 | -10.08 | -11.98 | -0.43 | 8.24 | 11.97 | 0.62 |
| OPTL | 11 | 58 | -8.84 | -6.28 | -0.39 | 5.88 | 6.07 | 0.85‡ |
| NONPA RA | 4 | 56 | -12.48 | -34.21 | 0.00 | 13.73 | 39.03 | 0.90† |
| MV | 7 | 47 | -9.27 | -12.30 | -0.38 | 8.04 | 10.55 | 0.86† |
| SVAR | 5 | 56 | -13.68 | -22.39 | -1.00† | 10.59 | 20.57 | 0.83† |
| UC | 5 | 50 | -10.13 | -20.08 | -1.00† | 10.00 | 20.50 | 1.00† |
| Averages | 5.90 | 55 | -11.28 | -20.34 | -0.57 | 10.08 | 23.74 | 0.79 |

† denotes significance of the concordance statistic at the five percent level.
‡ denotes significance of the concordance statistic at the one percent level.

**Table 2:**
**Correlation statistics**

|  | LIN | SEG | DIFF4 | BPF | HP1600 | OPTL | NON-PARA | MV | SVAR | UC |
|---|---|---|---|---|---|---|---|---|---|---|
| LIN | 1.00 | 0.98‡ | 0.37‡ | 0.81‡ | 0.82‡ | 0.71‡ | 0.92‡ | 0.76‡ | 0.56‡ | 0.85‡ |
| SEG |  | 1.00 | 0.35‡ | 0.83‡ | 0.85‡ | 0.74‡ | 0.94‡ | 0.75‡ | 0.65‡ | 0.87‡ |
| DIFF4 |  |  | 1.00 | 0.24‡ | 0.54‡ | 0.58‡ | 0.32‡ | 0.72‡ | 0.23† | 0.32‡ |
| BPF |  |  |  | 1.00 | 0.75‡ | 0.61‡ | 0.77‡ | 0.65‡ | 0.49‡ | 0.78‡ |
| HP1600 |  |  |  |  | 1.00 | 0.97‡ | 0.85‡ | 0.92‡ | 0.54‡ | 0.75‡ |
| OPTL |  |  |  |  |  | 1.00 | 0.77‡ | 0.89‡ | 0.46‡ | 0.63‡ |
| NONPARA |  |  |  |  |  |  | 1.00 | 0.68‡ | 0.53‡ | 0.77‡ |
| MV |  |  |  |  |  |  |  | 1.00 | 0.50‡ | 0.73‡ |
| SVAR |  |  |  |  |  |  |  |  | 1.00 | 0.79‡ |
| UC |  |  |  |  |  |  |  |  |  | 1.00 |

† denotes significance of the concordance statistic at the five percent level.
‡ denotes significance of the concordance statistic at the one percent level.

**Table 3:**
**Concordance statistics**

|  | LIN | SEG | DIFF4 | BPF | HP1600 | OPTL | NON-PARA | MV | SVAR | UC |
|---|---|---|---|---|---|---|---|---|---|---|
| LIN | 1.00 | 0.95‡ | 0.59 | 0.77‡ | 0.73‡ | 0.75‡ | 0.91‡ | 0.77‡ | 0.89‡ | 0.77‡ |
| SEG | | 1.00 | 0.61† | 0.81‡ | 0.75‡ | 0.75‡ | 0.95‡ | 0.77‡ | 0.89‡ | 0.79‡ |
| DIFF4 | | | 1.00 | 0.63 | 0.69‡ | 0.55‡ | 0.61† | 0.64‡ | 0.62‡ | 0.61† |
| BPF | | | | 1.00 | 0.78‡ | 0.70‡ | 0.77‡ | 0.68‡ | 0.64‡ | 0.79‡ |
| HP1600 | | | | | 1.00 | 0.92‡ | 0.86‡ | 0.79‡ | 0.72‡ | 0.82‡ |
| OPTL | | | | | | 1.00 | 0.78‡ | 0.80‡ | 0.75‡ | 0.77‡ |
| NONPARA | | | | | | | 1.00 | 0.76‡ | 0.72‡ | 0.87‡ |
| MV | | | | | | | | 1.00 | 0.67‡ | 0.78‡ |
| SVAR | | | | | | | | | 1.00 | 0.80‡ |
| UC | | | | | | | | | | 1.00 |

† denotes significance of the concordance statistic at the five percent level.
‡ denotes significance of the concordance statistic at the one percent level.

**Figure 1**
**Transfer functions**

First difference filter
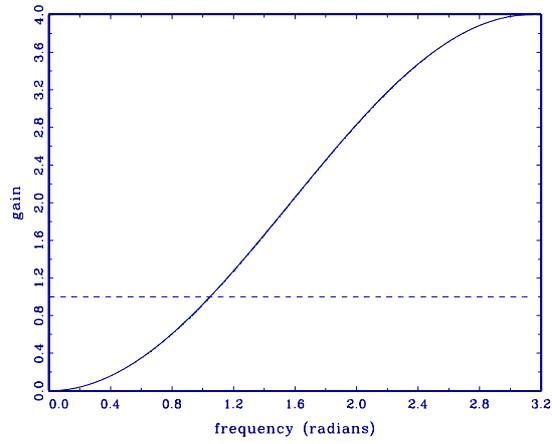


Fourth difference filter
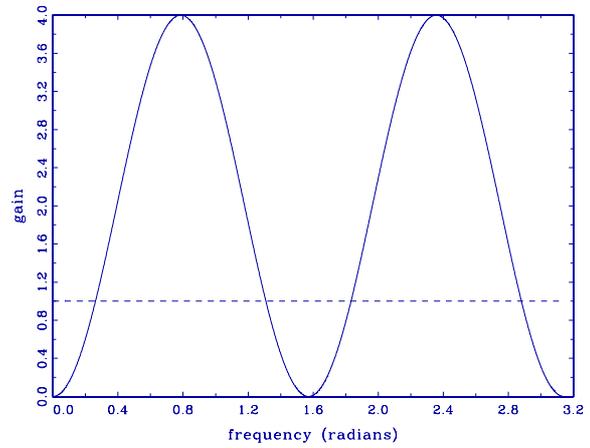


Hodrick-Prescott filter
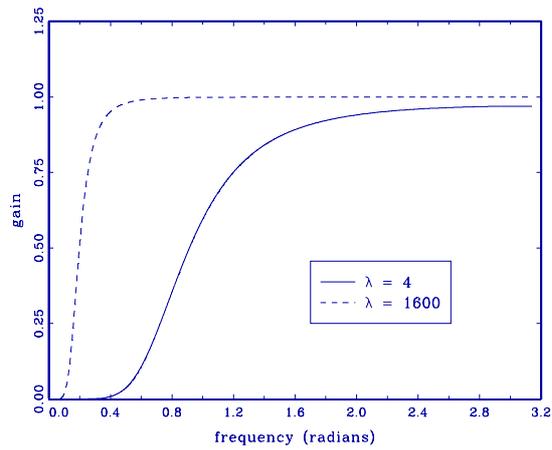


Band-pass filter
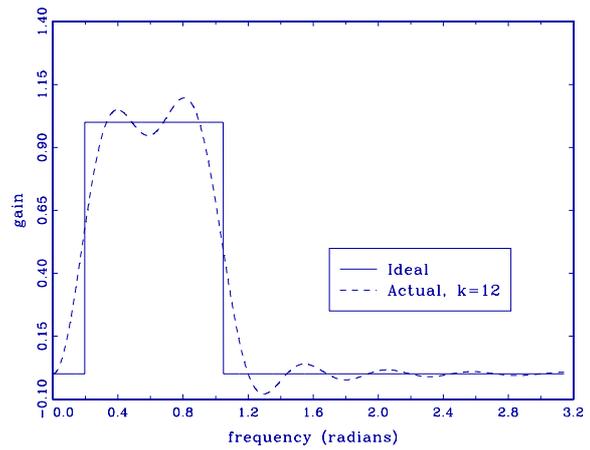
**Figure 2a**
**Detrending results**

Linear trend
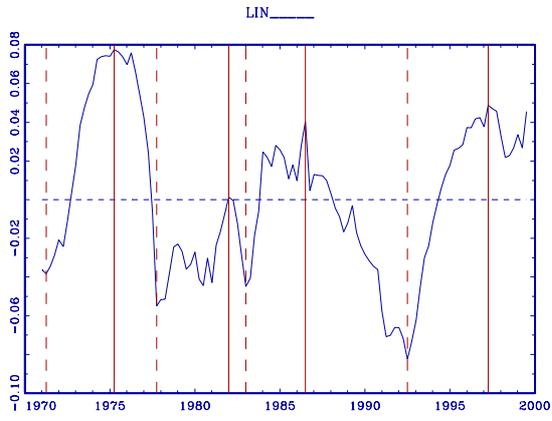


Segmented trend
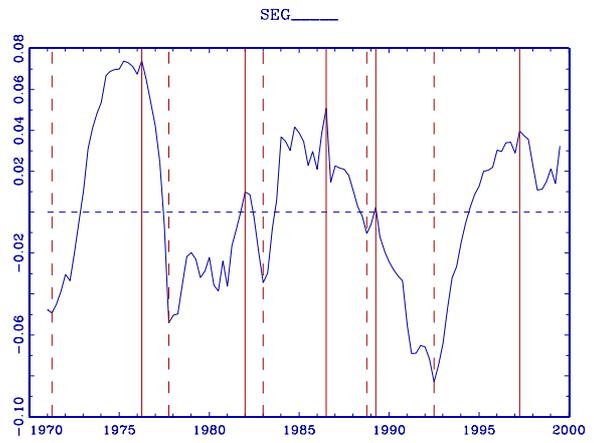


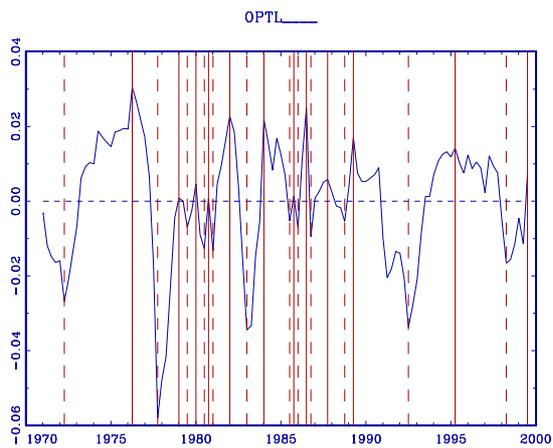Fourth difference filter
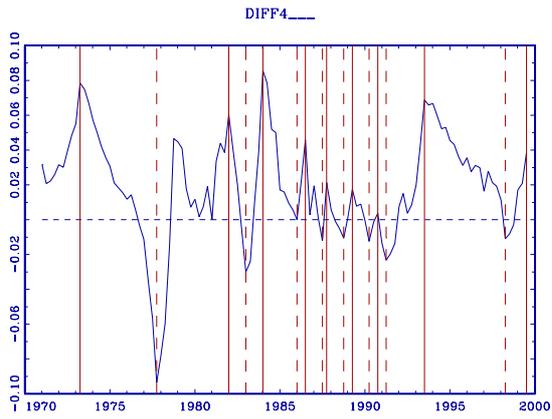
**Figure 2b**
**Detrending results**

Hodrick-Prescott filter, $\lambda = 1600$

HP1600__

Band-pass filter

BPF_____
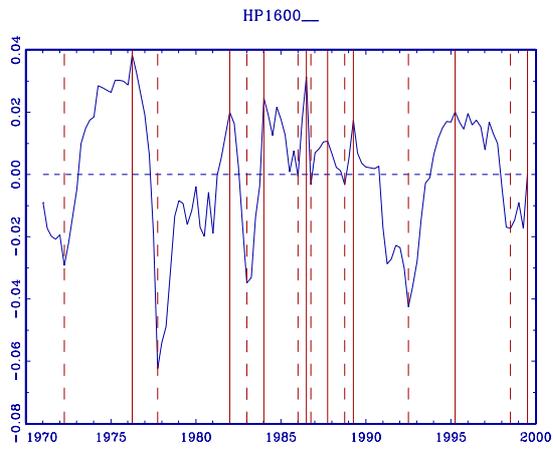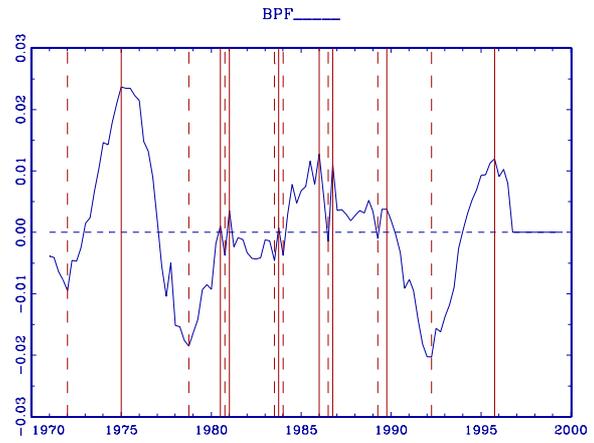
Optimal $\lambda$

OPTL____

Nonparametric estimator

NONPARA_
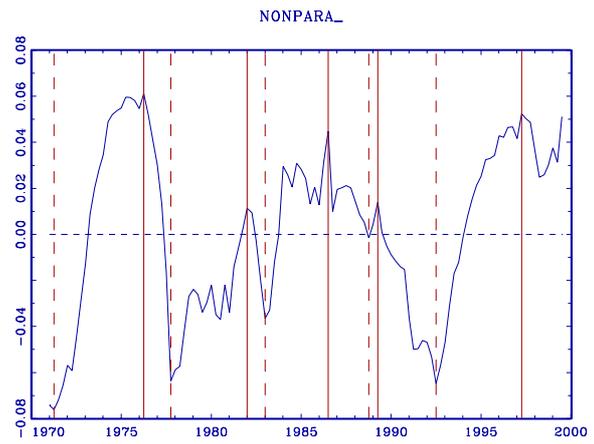
**Figure 3**
**Output gap measures**

MV filter
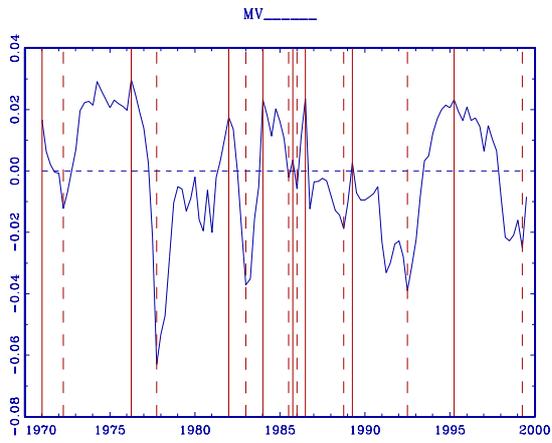
Structural VAR
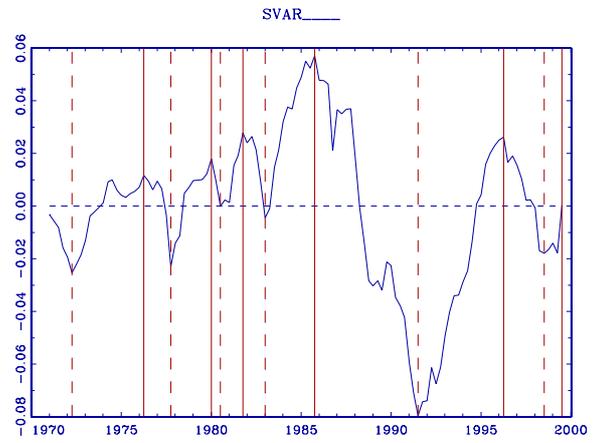


Multivariate unobserved components
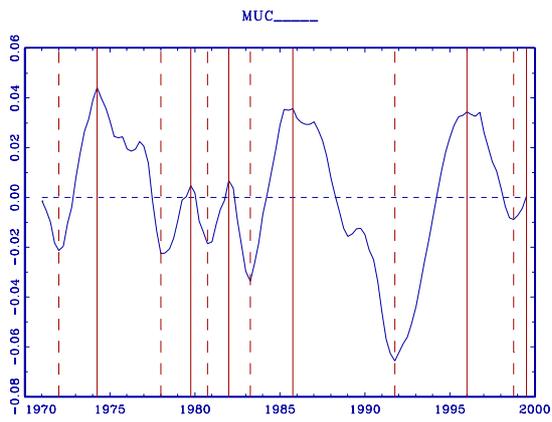
**Figure 4**
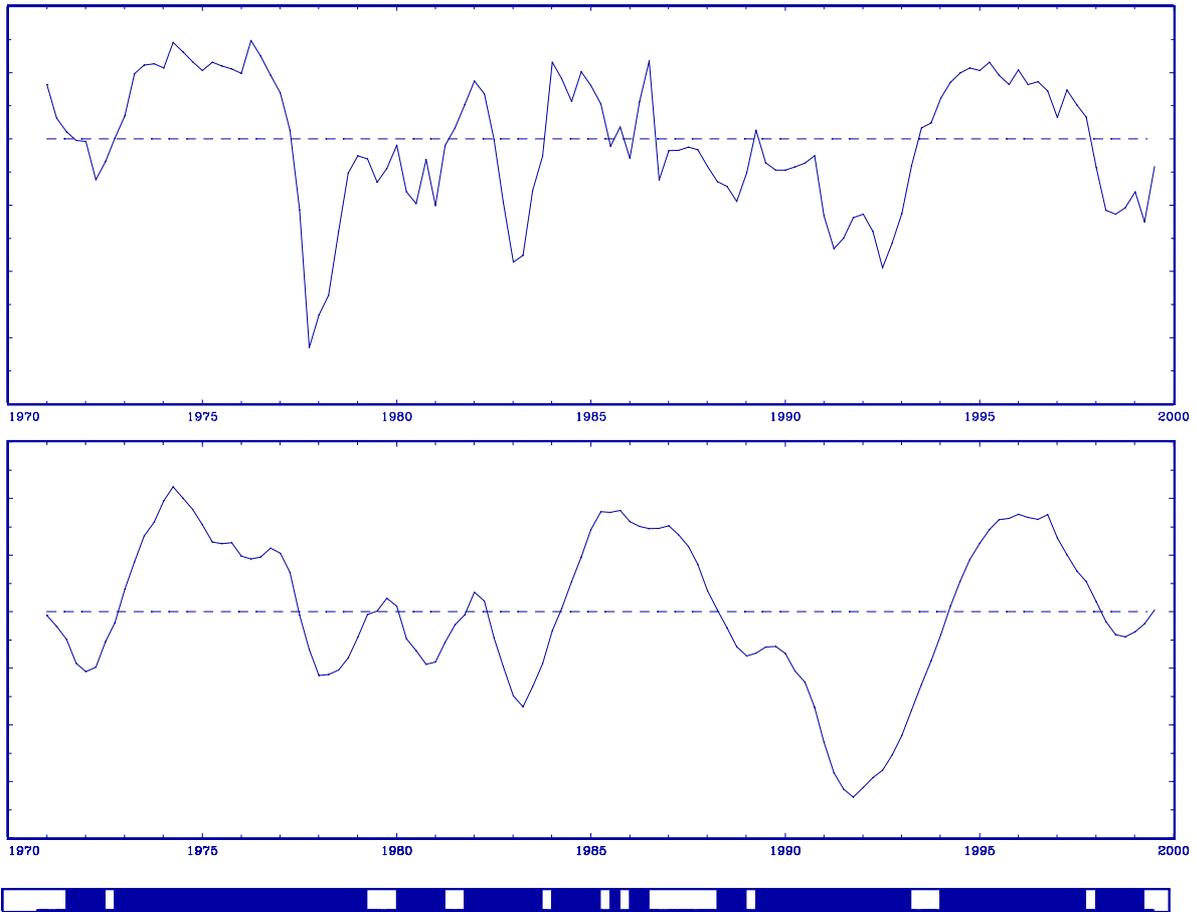**Phase behaviour of MV and UC gaps**
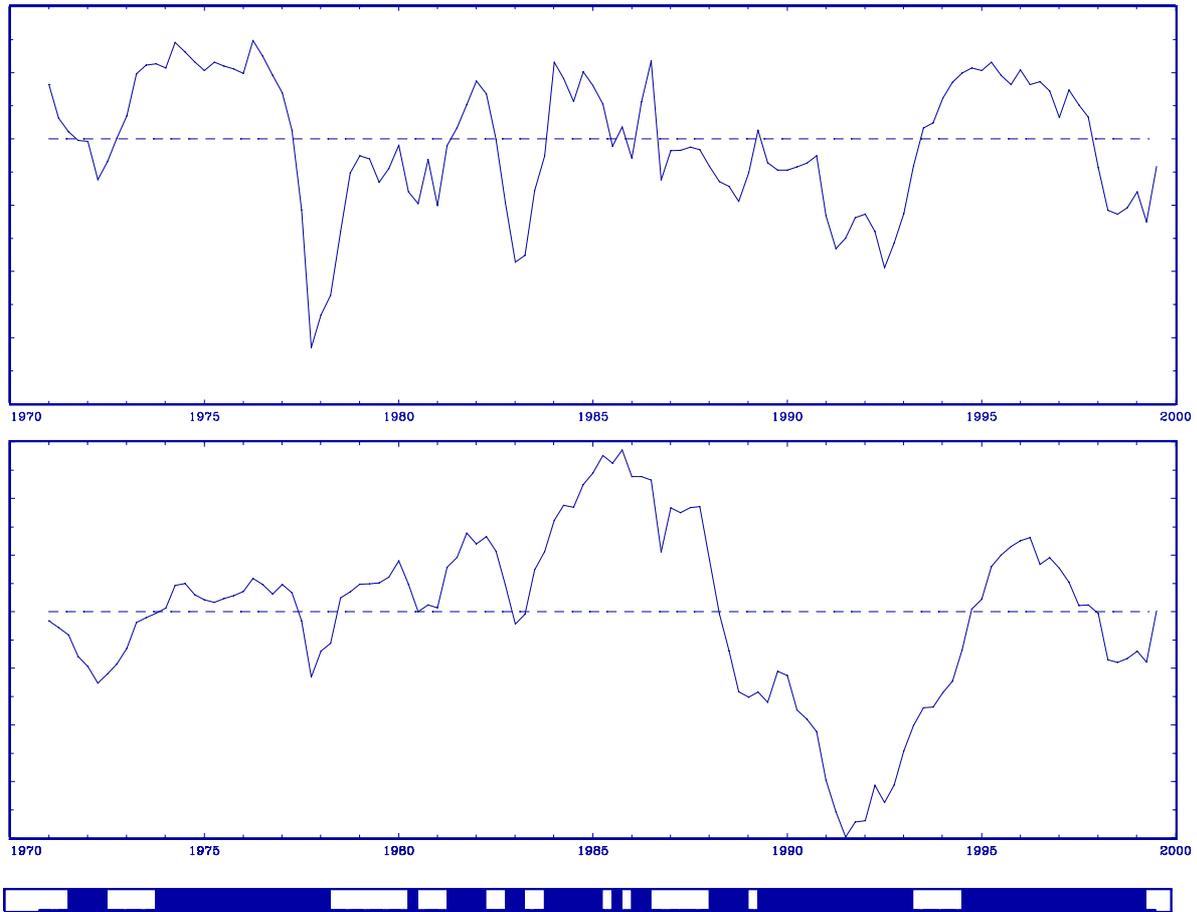
**Figure 5**
**Phase behaviour of MV and SVAR gaps**

**Figure 6**
**Phase behaviour of UC and SVAR gaps**