# Fables of Faubus?: Testing the Sectoral Shift Hypothesis in the Netherlands Using a Simplified Kalman Filter Model[1]

Ivo De Loo,
MERIT, Maastricht University

February 1998

## Abstract

The presence of structural breaks can seriously affect the outcome of standard regression methods like OLS. Although there are many methods available to deal with them, we focus here on a particular linear filtering method, namely the Kalman Filter. Its results vis a vis a regular OLS approach are illustrated by testing the sectoral shift hypothesis in the Netherlands. Although a rather simplified version of the Kalman Filter is used, it turns out to be a sufficient enough approximation. What we find, is that the variables capturing the sectoral shift hypothesis are the most important in explaining Dutch unemployment behaviour during the postwar period. Thus, the hypothesis is endorsed. On the other hand, our highly significant constant term indicates that the inclusion of other variables affecting unemployment may alter the results. Our conclusion thus is a tentative one.

Maastricht Economic Research Institute on Innovation and Technology, University of Maastricht, P.O. Box 616, 6200 MD Maastricht, The Netherlands, tel: +31 43 3883867, fax: +31 43 3216518, Email: i.deloo@merit.unimaas.nl

---

## 1. Introduction

Analyzing structural breaks in time series forms an important part of economic literature[1]. Also, there is a huge body of work in econometrics on this subject[2]. In the latter category, one can distinguish two approaches to deal with structural breaks. These approaches are:

1. Estimating a model as if no structural breaks were present and then testing whether this is correct (called *ex-post methods*);

2. Allowing for the presence of structural breaks right from the start by choosing an estimation method that takes (possibly) varying parameter values into account (called *ex-ante methods*).

The first approach is by far the most commonly used, and includes methods like the Chow-Test (or variations on it). However, the second approach is the most appealing, but has had little practical use in economic literature over the past few years. It is a particular subset of methods in the latter approach that will be discussed in this paper, namely the *Kalman Filter*.

After a general introduction to structural breaks, the Kalman Filter will be formally presented in the second Section. Then, in Section 3, it will be applied to a model explaining sectoral shifts in the Dutch unemployment rate during the postwar period. Finally, Section 4 comprises some general conclusions and guidelines for further research.

## 2. The Kalman Filter

What are 'structural breaks'? This is a question that cannot be answered immediately. A concept like a 'break' implies a degree of change, which needs measurement. From an economic point of view, these changes include shifts in the composition of output or input vectors. For example, (persistent) changes in (sectoral) imports and exports, in the composition of the labour force or GDP can be viewed as structural breaks. However, taking a statistical point of view, the notion of structural change is related to model structure, focusing both on the variables defining the structure and the parameters values associated with them.

Usually, when testing an economic model empirically, some form of Ordinary Least Squares (OLS) or Generalized Least Squares (GLS) estimation is applied to a predetermined (set of) equation(s). In doing so, we assume constant regression coefficients over time (by construction). However, if they do change, these techniques will yield incorrect conclusions with regard to the relationship being studied. Although there are methods that can detect whether the assumption of constant regression coefficients is valid, these are only applied afterwards. Methods one can think of include Recursive Least Squares (RLS), One-Step Residuals and (Scaled Recursive) Chow-Tests. Charemza and Deadman (1992) discuss these methods in detail.

---

1 See for example Maddison and van der Wee (1994) and Vilares (1986).

2 Like Kraemer (ed.) (1989) and Broemeling and Tsurumi (1987). In 1982, a complete issue of the *Journal of Econometrics* was dedicated to modelling structural breaks.

One the other hand, we could opt for modelling coefficient variability ex-ante. Then we would allow for stochastic regression coefficients right from the start, whereas they were fixed in the previous, ex-post context[3]. One of the methods available is the *Kalman Filter*, a linear filtering method introduced by Kalman (1960) and first used in economics by Rosenberg (1968). Its popularity grew steadily during the seventies and eighties[4], but since then, it has hardly been used in economics[5]. Any Kalman Filter consists of two parts, namely:

1. A *transition equation*, which describes the evolution of a set of unobservable, so-called *state variables* over time;

2. A *measurement equation*, which describes how well the actually observed data are generated from the state variables.

In our context, the state variables would be the regression coefficients to be estimated. Together, the transition and measurement equations constitute a *state space model*. A general Kalman Filter can be formulated as follows:

$$
\begin{cases}
\boldsymbol{y}_t = Z_t\alpha_t + S_t\delta_t \\
\alpha_t = \Phi_t\alpha_{t-1} + \boldsymbol{c}_t + \Psi_t\boldsymbol{v}_t
\end{cases}
\tag{2.1}
$$

where, $Z_t$ = (N$x$K) matrix;
$S_t$ = (N$x$M) matrix;
$\boldsymbol{y}_t$ = (N$x$1) vector;
$\alpha_t$ = (K$x$1) vector of state variables;
$\boldsymbol{c}_t$ = non-stochastic (K$x$1) vector;
$\Phi_t$, $\Psi_t$ = fixed (K$x$K) and (K$x$m) matrices respectively;
$\delta_t$, $\boldsymbol{v}_t$ = M- and m-dimensional white noise vectors respectively.

It is assumed that the $\delta_t$ and $\boldsymbol{v}_t$ terms are uncorrelated with both each other and the prior state vector $\alpha_0$ ($\forall\ t = 1,2,...,T$). Furthermore, although the variance of $\delta_t$ and $\boldsymbol{v}_t$ may alter over time, their elements are assumed not be correlated at different points in time.

The first equation in model (2.1) is an ordinary regression equation with Time-Varying Parameters, whereas the second equation describes the evolution of the parameters considered. The setup of the Kalman Filter is very broad, and includes all ARMA *(r,q)* models[6].

By means of the Kalman Filter, an estimator will be updated as soon as new observations become available. This process consists of two parts. First, an optimal prediction of the next observation is made, given all information currently available. Then, the new observation is included in the estimator of the state vector so that the estimates can be updated in the next round.

---

3 Although the estimated values of the regression coefficients were also stochastic previously.

4 For example, see Burmeister and Wall (1982), Conrad and Corrado (1979), Engle (1978) and Howrey (1978).

5 Though recent applications include Amable and Luillard (1997) and Hall *et al.* (1992).

6 See Harvey (1981), pp. 103-104 for details.

Usually, the data are assumed to follow a normal distribution. Yet, if this assumption is relaxed, the Kalman Filter will still produce an 'optimal' solution[7]. Given the recursive nature of the Kalman Filter in determining estimates of $\alpha_t$ (using new information on both $y_t$ and $Z_t$), it can be called a Bayesian method. Estimation takes place by means of Maximum Likelihood techniques[8].

The recursive nature of the Kalman Filter assures that shifts in variables over time will be picked up as soon as they occur, so that the corresponding shifts in parameter values will (only) be accounted for from that point in time onwards. Thus, the fit of a model estimated by means of a Kalman Filter should in general be at least comparable to, say, the fit generated by the OLS procedure[9].

Nevertheless, using techniques like the Kalman Filter to pick up structural breaks assumes that a 'genuine' break has occurred. Yet, it is extremely difficult to distinguish between model misspecification and structural breaks. Model specifications are irreversibly linked to structural changes and breaks, and their possible effects should therefore be part of any model structure.

Also, the data entering a model should be carefully checked a priori[10]. Changes in, say, the aggregation of sectors or the measurement of variables may point to 'structural' breaks that are not structural in nature at all: they are merely caused by alterations of definitions. Therefore, one should not assume structural breaks to be present too soon, certainly if no obvious economic reasons can be found for them.

In Section 3, structural changes in the Dutch unemployment rate between 1952-1993 will be analyzed using a simplified Kalman Filter setup. There, we will mostly take an economic point of view regarding structural change. That is to say, we will assume (at least, in principle) that the model structure is a reasonable enough approximation of reality.

## 3. Model and Results

Generally, it takes time before workers who become unemployed find a new job, even more when they change jobs between sectors (provided that they find one). This labour reallocation process gives rise to cyclical unemployment. But how is cyclical unemployment embedded in economic theory?

---

7 Although then the concept of *Minimum Mean Square (Linear) Estimators (MMSE)* is of importance, cf. Harvey (1981), pp. 104-105. Roughly speaking, the Kalman Filter assures that, independent of the distribution of the data underlying the model, the prediction of the estimation error between $\alpha_t$ and $a_t$ (the outcome of the model if $\alpha_t$ was fixed) is minimized.

8 See Harvey (1981), Ch. 4.4 for details.

9 The OLS procedure can be shown to be a special case of the Kalman Filter, cf. Watson (1983).

10 Although this is a topic often taken for granted.

On the one hand, we could say that cyclical unemployment is a deviation from a more or less stable natural rate. Alternatively, cyclical unemployment could be defined as a fluctuation of the natural rate itself. The latter perspective was introduced by Lilien (1982), and is called the *sectoral shift hypothesis.* Over the years, many papers were issued on mostly American data, showing evidence both in favour and against the hypothesis[11].

In this Section, we will test the Lilien hypothesis using Dutch data. We will do so by estimating a model with the Kalman Filter. This approach has the advantage that we can use time series published by the Central Planning Bureau (CPB) from 1950 onwards, although there have been revisions of the data in 1969, 1977 and 1985 (leading to breaks in the original series). Although the revisions have already been pulled back until 1977, using the Kalman Filter ensures that the breaks that occur because of this show up in the (residuals of the) transition equation[12].

Lilien proxied the dispersion of employment by means of a weighted standard deviation variable using sectoral employment data:

$$\hat{\sigma}_t = \sqrt{\sum_i (\frac{x_{it}}{\sum_i x_{it}})(\nabla \ln x_{it} - \nabla \ln \sum_i x_{it})^2} \; ,$$

where, $x_{it}$ = employment in sector $i$ ($i = 1,2,...,N$);
$\nabla$ = difference operator.

$\hat{\sigma}_t$ is calculated on the basis of the assumption that only sectoral shifts influence the dispersion of employment. However, employment dispersion can also be affected by aggregate shocks (which are also included in $\hat{\sigma}_t$). Therefore, Mills (1996) proposed a method to eliminate these influences in $\hat{\sigma}_t$. His method can be described as follows.

Suppose we have a regression model that contains $\hat{\sigma}_t$ (and possibly, its lagged values) among the explanatory variables. Each relative sectoral growth rate $\nabla x_{it} - \nabla(\sum_i x_{it})$ is then regressed on its own lagged value and current and lagged values of other variables that influence unemployment (and which may or may not be a part of the model already). The residuals of this regression, $z_{it}$, are then used to construct the revised dispersion measure $\hat{\sigma}_t^p$:

$$\hat{\sigma}_t^p = \sqrt{\sum_i (\frac{x_{it}}{\sum_i x_{it}}) z_{it}^2} \qquad\qquad (3.1)$$

Formula (3.1) will enter our analysis as a measure of Lilien's sectoral shift hypothesis. Following Barro (1977) and Mills (1996), we will try to explain the Dutch unemployment rate as a function of its own lagged values, money growth, interest rate changes and the dispersion measure depicted

---

11 See Mills (1996) for an overview. A more recent publication is Fortin and Araar (1997).

12 See also note 19.

above. This implies that we will mainly use data stemming from a monetary world. The interest and money variables are included to capture the hypothesis that inflation and monetary policies, if unexpected, affect unemployment duration. During the modelling process, the *Hendry/LSE (a.k.a. general to specific modelling) approach*[13] will be adopted. This means that we will start with a very broad model, which will gradually be reduced to reach a 'final', smaller model. The following regression model will be initially estimated[14]:

$$
\left\{ u_t = \begin{bmatrix} \beta_{0t} \\ \beta_{1t} \\ \beta_{2t} \\ \beta_{3t} \end{bmatrix} \begin{bmatrix} 1 \\ u_{t-1} \\ u_{t-2} \\ u_{t-3} \end{bmatrix} + \begin{bmatrix} \gamma_{0t} \\ \gamma_{1t} \\ \gamma_{2t} \\ \gamma_{3t} \end{bmatrix} \begin{bmatrix} M_t \\ M_{t-1} \\ M_{t-2} \\ M_{t-3} \end{bmatrix} + \begin{bmatrix} \delta_{0t} \\ \delta_{1t} \\ \delta_{2t} \\ \delta_{3t} \end{bmatrix} \begin{bmatrix} r_t \\ r_{t-1} \\ r_{t-2} \\ r_{t-3} \end{bmatrix} + \begin{bmatrix} \eta_{0t} \\ \eta_{1t} \\ \eta_{2t} \\ \eta_{3t} \end{bmatrix} \begin{bmatrix} \hat{\sigma}_t^p \\ \hat{\sigma}_{t-1}^p \\ \hat{\sigma}_{t-2}^p \\ \hat{\sigma}_{t-3}^p \end{bmatrix} + \varepsilon_t \\[3em] \begin{bmatrix} \beta_{it} \\ \gamma_{it} \\ \delta_{it} \\ \eta_{it} \end{bmatrix} = \begin{bmatrix} \beta_{i,t-1} \\ \gamma_{i,t-1} \\ \delta_{i,t-1} \\ \eta_{i,t-1} \end{bmatrix} + \begin{bmatrix} v_{\beta_{it}} \\ v_{\gamma_{it}} \\ v_{\delta_{it}} \\ v_{\eta_{it}} \end{bmatrix} \right. \qquad (3.2, )
$$

where, $u_t$ = (national) unemployment rate;
$M_t$ = money supply ($M_1$)[15];
$r_t$ = short-term interest rate;
$\varepsilon_t$ = random disturbance term;
$i$ = 0,1,2,3[16].

We expect the interest rate variable to be negatively related to unemployment, whereas the opposite is assumed for the money variable. All other variables can have either sign (although, leaving aside the constant term, positive signs are the most likely).

In order to construct $\hat{\sigma}_t^p$, $\hat{\sigma}_t$ is regressed[17] on its own lagged value and current and (once) lagged values of both $M_t$ and $r_t$ (20 sectors are used in the analysis). Also, the current value of the GDP price level (with 1985 as a base year) is included in the regression to capture the influence of national

---

13 Cf. Hendry (1985).

14 As compared to model (2.1), we have set $S_t$, $\Phi_t$ and $\Psi_t$ equal to identity matrices, and put all elements of $c_t$ equal to zero (for reasons of simplicity). Furthermore, if the disturbance term in the transition equation of model (3.2), $v_{jit}$, were dropped, a RLS model would have resulted. Thus in fact, we will be using a somewhat 'extended' RLS model. See also note 28. Information about data sources is given in Appendix A.

15 A narrow definition of money is used, cf. Pelloni (1992).

16 All variables are lagged three times following the work by Garcia-Ferrer *et al.* (1987).

17 Details about its calculation can be found in Appendix A.

commodity prices. The residuals of this regression, estimated by OLS, enter the calculation of $\hat{\sigma}_t^p$ according to model (3.1).

In addition, the residuals of the measurement equation of model (3.2), $\varepsilon_t$, are examined on the presence of a unit root. When a time series contains one (or more) structural break(s), applying standard unit root tests like Dickey-Fuller could erroneously lead to the conclusion that the series contains a unit root, whereas in fact it does not. Although alternative measures have been proposed[18], it is here that the features of the Kalman Filter come into view. Since it will pick up structural breaks by construction[19], there is no need for applying alternative measures when testing for a unit root: the standard Dickey-Fuller test will still be valid. However, if we estimate the first equation of model (3.2) by OLS, then we would have to resort to alternative measures. As an alternative to Dickey-Fuller, the Phillips-Perron test will be used (which is a variation on Dickey-Fuller, able to deal with structural breaks)[20].

If a unit root turns out to be present in the (recursive) Kalman Filter residuals anyway, we have to opt for expressing model (3.2) in first differences: $u_t$ then becomes $\triangledown u_t$, and the first equation of model (3.2) would have to be rearranged accordingly.

To test the sectoral shift hypothesis, we gradually reduce the size of model (3.2), dropping insignificant variables one at a time, while testing for the (statistical) acceptability hereof by means of Likelihood Ratio tests[21]. We also try to eliminate variables with incorrect signs, and delete them even if the $t$-test says that the coefficients do differ significantly from zero (as long as we can conclude from the Likelihood Ratio test that they can be dropped together with all other previously deleted variables). A 5% level of significance is used. In any case, the constant term will not be dropped, even if it turns out to be insignificantly different from zero, for it measures the mean effect of possibly missing variables in our model[22]. The outcome of the two aforementioned unit root tests is also shown. There, the values between brackets indicate $p$-values.

The 'final' model we end up with is, together with the corresponding OLS estimates, shown in Table 3.1 below. Because the observations over the period 1955-1961 are necessary to estimate the prior state vectors $\beta_{i0}$, $\gamma_{i0}$, $\delta_{i0}$

---

18 See Amsler and Lee (1995) for (both an addition and) an overview.

19 As long as a correct filtering scheme has been chosen, the effects of structural breaks are taken up by the coefficients themselves, and not by the residuals of either the state or the transition equation.

20 Cf. Phillips and Perron (1988).

21 Tests based on a regular Wald statistic were carried out too, leading to identical results in the end.

22 The approach that is chosen to model unemployment, namely based on monetary data, is not the only one that has attracted much attention. The constant term intends, in a way, to capture some of the effects of all other theories in this field (like the production function view described in Hamermesh and Grant (1979) and Pasinetti's (1981) view, stressing the learning power of individuals in a community).

and $\eta_{i0}$ (which thus remain the same until 1962), the OLS estimates are confined to the period 1962-1993. In this way, the comparability of the results is assured.

**Table 3.1**  OLS and Kalman Filter estimates of the 'final' version of model (3.2) (either *t*-values or *p*-values are shown between brackets. Coefficients that do not appear in the Table have been put equal to zero).

| Estimation method<br><br>Coefficient | OLS | Kalman Filter | Unit root tests[23] | |
|---|---|---|---|---|
| $\beta_0$ | -.245 (-.61) | 1.019 (2.95) | **OLS (#lags=2):** | |
| $\beta_1$ | 1.789 (11.5) | .481 (1.84) | Dickey-Fuller | -2.500 (.33) |
| $\beta_2$ | -1.159 (-5.69) | -.685 (-2.77) | Phillips-Perron | -30.25 (.01) |
| $\gamma_1$ | .037 (3.76) | .093 (3.76) | | |
| $\delta_0$ | -.035 (-.56) | -.251 (-1.12) | **Kalman Filter (#lags=5):** | |
| $\eta_1$ | 26.97 (2.31) | 101.1 (14.6) | Dickey-Fuller | -3.856 (.01) |
| $\eta_2$ | 23.25 (2.43) | 40.92 (4.07) | Phillips-Perron | -34.78 (.00) |
| $R^2_{adj.}$ | .973 | 1.000 | **LR statistic**[24] | -33.12 (1.00) |

It may be illuminating to see how the importance of the variables has changed over time. Therefore, Table 3.2 contains the evolving state vectors of the 'final' Kalman Filter model.

---

23 The optimal lag length used in the unit root tests is calculated by means of a variation on the AIC (Akaike Information Criterion), described in Pantula *et al.* (1994).

24 As compared to the Kalman Filter estimates of the 'full' model.

**Table 3.2**     Evolving state vectors of the 'final' model (continued on next page).

| Year | $\beta_0$ | $\beta_1$ | $\beta_2$ | $\delta_0$ | $\gamma_1$ | $\eta_1$ | $\eta_2$ |
|------|-------|-------|--------|--------|--------|-------|-------|
| 1955 | 1.194 | 0.511 | -0.724 | -0.437 | 0.020 | 116.4 | 35.28 |
| 1956 | 1.194 | 0.511 | -0.724 | -0.437 | 0.020 | 116.4 | 35.28 |
| 1957 | 1.194 | 0.511 | -0.724 | -0.437 | 0.020 | 116.4 | 35.28 |
| 1958 | 1.194 | 0.511 | -0.724 | -0.437 | 0.020 | 116.4 | 35.28 |
| 1959 | 1.194 | 0.511 | -0.724 | -0.437 | 0.020 | 116.4 | 35.28 |
| 1960 | 1.194 | 0.511 | -0.724 | -0.437 | 0.020 | 116.4 | 35.28 |
| 1961 | 1.194 | 0.511 | -0.724 | -0.437 | 0.020 | 116.4 | 35.28 |
| 1962 | 1.123 | 0.475 | -0.716 | -0.376 | 0.047 | 115.1 | 38.39 |
| 1963 | 1.145 | 0.485 | -0.718 | -0.396 | 0.028 | 115.3 | 37.31 |
| 1964 | 1.010 | 0.418 | -0.706 | -0.292 | 0.032 | 113.7 | 43.42 |
| 1965 | 1.001 | 0.414 | -0.704 | -0.282 | 0.043 | 113.6 | 43.83 |
| 1966 | 0.880 | 0.355 | -0.694 | -0.180 | 0.054 | 112.0 | 49.20 |
| 1967 | 0.955 | 0.397 | -0.699 | -0.229 | 0.093 | 113.0 | 45.84 |
| 1968 | 0.932 | 0.381 | -0.698 | -0.214 | 0.079 | 112.8 | 46.90 |
| 1969 | 0.814 | 0.322 | -0.674 | -0.105 | 0.096 | 111.2 | 51.98 |
| 1970 | 0.742 | 0.289 | -0.658 | -0.048 | -0.108 | 102.3 | 54.12 |
| 1971 | 0.731 | 0.276 | -0.665 | -0.095 | -0.029 | 101.4 | 55.49 |
| 1972 | 1.073 | 0.441 | -0.685 | -0.346 | 0.081 | 104.2 | 38.95 |
| 1973 | 1.065 | 0.440 | -0.685 | -0.333 | 0.088 | 104.1 | 39.29 |
| 1974 | 1.015 | 0.423 | -0.655 | -0.263 | 0.132 | 103.2 | 41.40 |
| 1975 | 0.998 | 0.412 | -0.656 | -0.241 | 0.116 | 102.9 | 42.13 |
| 1976 | 0.991 | 0.415 | -0.656 | -0.228 | 0.134 | 102.7 | 42.39 |
| 1977 | 0.980 | 0.405 | -0.662 | -0.215 | 0.119 | 102.5 | 42.88 |
| 1978 | 0.982 | 0.406 | -0.663 | -0.216 | 0.116 | 102.5 | 42.82 |
| 1979 | 0.980 | 0.405 | -0.663 | -0.213 | 0.121 | 102.5 | 42.90 |
| 1980 | 0.981 | 0.404 | -0.663 | -0.216 | 0.111 | 102.5 | 42.87 |
| 1981 | 0.979 | 0.404 | -0.663 | -0.213 | 0.122 | 102.5 | 42.97 |
| 1982 | 0.996 | 0.435 | -0.658 | -0.234 | 0.161 | 102.6 | 42.24 |
| 1983 | 1.007 | 0.454 | -0.654 | -0.250 | 0.181 | 102.7 | 41.66 |
| 1984 | 1.006 | 0.452 | -0.655 | -0.250 | 0.178 | 102.7 | 41.69 |
| 1985 | 1.010 | 0.453 | -0.674 | -0.253 | 0.154 | 102.6 | 41.53 |
| 1986 | 1.036 | 0.494 | -0.675 | -0.248 | 0.072 | 101.6 | 40.05 |
| 1987 | 1.016 | 0.474 | -0.688 | -0.246 | 0.114 | 101.0 | 41.10 |
| 1988 | 1.017 | 0.475 | -0.689 | -0.246 | 0.116 | 101.0 | 41.02 |

**Table 3.2**  Evolving state vectors of the 'final' model (continued).

| Year | $\beta_0$ | $\beta_1$ | $\beta_2$ | $\delta_0$ | $\gamma_1$ | $\eta_1$ | $\eta_2$ |
|---|---|---|---|---|---|---|---|
| 1989 | 1.018 | 0.476 | -0.689 | -0.248 | 0.113 | 101.1 | 40.98 |
| 1990 | 1.019 | 0.478 | -0.688 | -0.250 | 0.105 | 101.1 | 40.96 |
| 1991 | 1.019 | 0.480 | -0.686 | -0.251 | 0.097 | 101.1 | 40.93 |
| 1992 | 1.019 | 0.481 | -0.685 | -0.251 | 0.092 | 101.1 | 40.93 |
| 1993 | 1.019 | 0.481 | -0.685 | -0.252 | 0.093 | 101.1 | 40.92 |

Three coefficients are, at the end of the estimation period, smaller than their initial values (namely, $\beta_0$, $\beta_1$ and $\eta_1$), whereas the opposite happens for the remaining ones ($\beta_2$, $\delta_0$, $\gamma_1$ and $\eta_2$). Around 1970, all coefficients (except $\beta_2$) undergo rather large shocks as far as their magnitude is concerned. Most prominent perhaps, are the negative values that occur for $\gamma_1$ in 1970 and 1971. At the same time, $\beta_0$, $\beta_1$, $\delta_0$, $\eta_1$ and $\eta_2$ are either greatly increased or decreased. For some time in the mid 1970s and the beginning of the 1980s, $\beta_0$ falls below one once more. By the mid 1980s, all coefficients (except $\gamma_1$) have already moved towards their final values.

What all this evidence sums up to, is that there appears to be a lot of volatility in the data (and thus, the size of the coefficients) of the 'final' model, which may very well point to structural breaks in the structure of Dutch unemployment during the estimation period[25]. OLS, being more 'static' in nature by construction, fails to pick up these effects. Partly, the evolution of the Dutch unemployment rate is self-enforcing: two autoregressive terms turn up significantly in the 'final' model. It is interesting that these effects counteract with one another. Even stronger, the total effect is negative (leaving aside time subscripts)[26]. However, from this result we cannot conclude that in the long run, unemployment in the Netherlands tends to decrease all by itself: the other variables play a major role in the debate too. Besides, the significance of the constant term (being the fourth important variable in the model based on the *t*-statistics) points to possibly missing variables (for example, more macroeconomic oriented variables like wages or variables on union power)[27]. On the contrary, the OLS estimates would imply that no additional explanatory variables should be included in our model (indicated by the insignificant constant term).

What is clear in any case, is that the most important variables (judging from the *t*-ratios) are the lagged employment dispersion measures. Thus, the sectoral shift hypothesis is endorsed. Moreover, given the evolution of the $\eta_i$ variables over time, the role played by the hypothesis is definitely positive. As far as the size of its relationship with the unemployment rate is concerned, we see that it is much stronger in our approach than when applying OLS.

---

25 Admittedly, most of the breaks that have occurred are due to the previously mentioned data revisions.

26 Which is appropriate if we were studying the long-term implications of our model.

27 See also note 22.

One may recall that we have used a rather simplified version of the Kalman Filter to test the sectoral shift hypothesis. This may give rise to the question whether the filtering mechanism that was employed is 'good enough' for the current application[28]. Watson (1983) proposes a small sample test to answer this question, based on the properties of the one-step ahead prediction errors (to be called $\pi_t$). They should be normally and independently distributed with mean zero and a variance equal to one. If the independence property is violated, the version of the Kalman Filter chosen is not optimal.

Watson applies a Von Neumann Ratio test to analyze the property. Alternatively, we could resort to a Kolmogorov-Smirnov test on the standard normality of the prediction errors. In doing so, we can use two samples: one based on the entire estimation period (1955-1993) and one based on the period after the prior state vectors have been estimated (1962-1993). This last period is sometimes referred to as the period over which *pure Kalman Filtering* is performed[29]. Although for this reason, the latter period is to be preferred, we will present the outcome for both samples. Table 3.3 lists the relevant statistics.

**Table 3.3**       Outcome of the Kolmogorov-Smirnov test on the standard normality of the one-step ahead prediction errors of the 'final' model (*p*-values between brackets).

| Sample period | Kolmogorov-Smirnov test statistic (*p*-value) |
|---|---|
| 1955-1993 | .881 (.42) |
| 1962-1993 | .708 (.70) |

The Table above tells us that the prediction errors $\pi_t$ indeed follow a standard normal distribution (a 5% level of significance is applied). Yet, hereby we have not tested the independence requirement. This is achieved by calculating Pearson's correlation coefficient between $\pi_t$ and $\pi_{t-1}$ (which is a valid approach for the normality of the data has been assessed)[30] over the two sample periods[31]. The results hereof are given in Table 3.4.

---

28 This question basically boils down to assessing whether the signal-to-noise ratio of this specific Kalman Filter when setting $S_t$, $\Phi_t$ and $\Psi_t$ equal to identity matrices, and all elements of $c_t$ equal to zero can (and should) be improved or not. A description of the underlying theory can be found in Anderson and Moore (1979), Ch. 5.5. However, we will use a slightly different method than the authors describe here.

29 See Watson (1983), p. 79.

30 Spearman rank correlation coefficients were also calculated, leading to identical results.

31 We only test for first-order autoregressive behaviour of the prediction errors. Both Pearson and Spearman correlation coefficients for lags two through five were calculated also, yielding similar (insignificant) results.

**Table 3.4**  Pearson correlation coefficient between the one-step ahead prediction errors of the 'final' model at period $t$ and period $t-1$ ($p$-values between brackets).

| Sample period | Pearson correlation coefficient ($p$-value) |
|---|---|
| 1955-1993 | -.086 (.61) |
| 1962-1993 | -.086 (.64) |

Looking at Table 3.4, we see that the independence assumption of $\pi_t$ is also satisfied. At a 5% level of significance we cannot reject the null hypothesis of zero correlation between $\pi_t$ and $\pi_{t-1}$ in either case. Nevertheless, an option that we still have not examined is that the one-step ahead prediction errors are biased. This may be the case because the basic assumptions underlying the Kalman Filter are highly restrictive: not only do the usual OLS assumptions have to be satisfied, but also should, among others, the initial state vectors $\beta_{i0}$, $\gamma_{i0}$, $\delta_{i0}$ and $\eta_{i0}$ be known in advance. Since we do not know them, and OLS regressions are carried out to estimate their initial values, a certain amount of bias could have slipped through in our final results[32]. To test for this notion, we can use a standard $t$-test on the mean of the one-step ahead prediction errors $\pi_t$[33]:

$$\Theta = \frac{\bar{\pi}}{\hat{\sigma}_{\bar{\pi}}} \sim t_{N-1},$$

where $\bar{\pi} = (1/N) \cdot \Sigma_i \pi_i$ ($i = 1,2,...,N$).

Again, we have a choice between the two aforementioned sample periods. We will present the results achieved with both, as shown in Table 3.5.

**Table 3.5**  Testing the bias of the one-step ahead prediction errors $\pi_t$ ($p$-values between brackets).

| Sample period | $\Theta$ ($p$-value) |
|---|---|
| 1955-1993 | .114 (.91) |
| 1962-1993 | .113 (.91) |

Our conclusion is independent of the fact which sample period we choose: in both cases, there seems to be no bias around the mean of the

---

32 For this reason, one may claim that it would have been better to use the Kalman Smoother instead of the Kalman Filter right from the start. Then, we would estimate our model by means of the Kalman Filter for period 1 through $N$, and given these estimates, from period $N$ to 1 and then again the other way around until the difference in estimates reached at stage $i$ do not differ substantially from those at stage $i+1$ (based on some criterion value). The estimates thus obtained are the steady state values of the model under consideration. However, the following test assesses whether these steady state values are sufficiently approximated by applying the Kalman Filter only once. If they turn out not to be, we will have to turn to using the Kalman Smoother indeed.

33 Following Watson (1983), p. 78.

prediction errors (using a 5% level of significance). Thus, although we opted for a rather simple Kalman Filter to analyze the sectoral shift hypothesis in the Netherlands, the previously obtained results are statistically valid. In short, sectoral shifts do seem to matter for the evolution of the Dutch unemployment rate. This result implies that policy makers should not limit themselves to aggregate models that do not explicitly incorporate a multisectoral dimension of (production and) employment when studying its behaviour. The significant constant term in our model also indicates that non-monetary and non-fiscal policies can play a substantial role in stabilizing unemployment in the Netherlands.

## 4. Summary and Conclusions

Regression results can be heavily influenced by the presence of structural breaks. Although there are many methods available to deal with them, we adopt a Kalman Filter approach here. Since we want to test the sectoral shift hypothesis in the Netherlands (stating that variations in the speed of labour reallocation between sectors results in fluctuations of the natural rate of unemployment), and the Dutch data have undergone several revisions, this approach seems to be more suitable than, say, running standard OLS regressions. Although we use a rather simplified version of the Kalman Filter to test the hypothesis, it turns out to be a statistically valid approach: the prediction errors are normally distributed and show no sign of bias. The sectoral shift hypothesis is tested within a framework of mostly monetary data, following the work of Barro (1977), Lilien (1982) and Mills (1996). What we find, is that the employment dispersion measures are the most important variables in explaining the Dutch unemployment behaviour during the postwar period. Thus, the sectoral shift hypothesis is endorsed. But, given our highly significant constant term, it may very well be that the inclusion of other variables affecting unemployment (for example, more macroeconomic oriented variables like wages or variables on union power), would alter our results somewhat. However, this must be left for further research. It is clear in any case that the nature of unemployment should be studied within a multisectoral framework, and that there obviously is room for more than just monetary and fiscal policies to stabilize it.

## References

Amable, B. and Juillard, M., 'The Historical Process of Convergence', *TSER Working Paper*, January 1997.

Amsler, C. and Lee, J., 'An LM Test for a Unit Root in the Presence of a Structural Change', *Econometric Theory* 11 (1995), pp. 359-368.

Anderson, B.D.O. and Moore, J.B., *Optimal Filtering*, Prentice Hall Incorporated, Englewood Cliffs, 1979.

Barro, R.J., 'Unanticipated Money Growth and Unemployment in the United States', *American Economic Review* 62 (1977), pp. 101-115.

Broemeling, L.D. and Tsurumi, H., *Econometrics and Structural Change*, Marcel Dekker Incorporated, New York, 1987.

Burmeister, E. and Wall, K.D., 'Kalman Filtering Estimation of Unobserved Rational Expectations with an Application to the German Hyperinflation', *Journal of Econometrics* 20 (1982), pp. 255-284.

Charemza, W.W. and Deadman, D.F., *New Directions in Economic Practice*, Edward Elgar Publishing Limited, Aldershot, 1992.

Conrad, W. and Corrado, C., 'Application of the Kalman Filter to Revisions in Monthly Retail Sales Estimates', *Journal of Economic Dynamics and Control* 1 (1979), pp. 177-198.

Engle, R.F., 'Estimating Structural Models of Seasonality'. In: Zellner, A. (ed.), *Seasonal Analysis of Economic Time Series*, Bureau of the Census, Washington, 1978.

Garcia-Ferrer, A. *et al.*, 'Macroeconomic Forecasting Using Pooled International Data', *Journal of Business & Economic Statistics* 5 (1987), pp. 53-67.

Fortin, M. and Araar, A., 'Sectoral Shifts, Stock Market Dispersion and Unemployment in Canada', *Applied Economics* 29 (1997), pp. 829-839.

Hall, S., Robertson, D. and Wickens, M., 'Measuring Convergence of the EC Economies', *London Business School Discussion Paper DP 1-92*, 1992.

Hamermesh, D.S., and Grant, J.H., 'Econometric Studies of Labour-Labour Substitution and Their Implications for Policy', *Journal of Human Resources* 14 (1979), pp. 518-542.

Harvey, A.C., *Time Series Models*, Phillip Allan, Oxford, 1981.

Hendry, D., 'Econometric Methodology: a Personal Perspective', 1985 (Paper Presented at the Econometric Society's Fifth World Congress, MIT).

Howrey, E.P., 'The Use of Preliminary Data in Econometric Forecasting', *Review of Economics and Statistics* 60 (1978), pp. 306-321.

Kalman, R.E., 'A New Approach to Linear Filtering and Prediction Problems', *Journal of Basic Engineering* 82 (1960), pp. 35-45.

Kraemer, W. (ed.), *The Econometrics of Structural Change*, Springer-Verlag, New York, 1989.

*International Financial Statistics Yearbook 1990*, Publication of the IMF, International Monetary Fund Publication Services, Washington, 1990.

*International Financial Statistics Yearbook 1995*, Publication of the IMF, International Monetary Fund Publication Services, Washington, 1995.

*LangeReeksen Boek 1950-1996*, Publication of the CPB, Afdeling Bedrijfstakkencoördinatie, The Hague, 1995.

Lilien, D.M., 'Sectoral Shifts and Cyclical Unemployment', *Journal of Political Economy* 90 (1982), pp. 777-793.

Maddison, A. and van der Wee, H., *Economic Growth and Structural Change: Comparative Approaches over the Long Run*, Bocconi University, Milan, 1994.

Mills, T.C., Pelloni, G. and Zervoyianni, A., 'Cyclical Unemployment and Sectoral Shifts: Further Tests of the Lilien Hypothesis for the UK', *Economics Letters* 52 (1996), pp. 55-60.

Pantula, S.G., Gonzalez-Farias, G. and Fuller, W.A., 'A Comparison of Unit Root Test Criteria', *Journal of Business & Economic Statistics* 12 (1994), pp. 449-459.

Pasinetti, L.L., *Structural Change and Economic Growth*, Cambridge University Press, Cambridge, 1981.

Pelloni, G., 'Sectoral Shifts and Unemployment Fluctuations in the UK', IDSE, *Dynamis Quaderno* 9/92, 1992.

Phillips, P.C.B. and Perron, P., 'Testing for a Unit Root in Time Series Regression', *Biometrika* 75 (1988), pp. 335-346.

Rosenberg, B., 'An Atomistic Approach to Linear Regression', *Berkeley Institute of Business and Economic Research Working Paper IP-137*, 1968.

Vilares, M.J., *Structural Change in Macroeconomic Models: Theory and Estimation*, Nijhoff, Dordrecht, 1986.

Watson, P.K., 'Kalman Filtering as an Alternative to Ordinary Least Squares - Some Theoretical Considerations and Empirical Results', *Empirical Economics* 8 (1983), pp. 71-85.

**Appendix A: Data Description**

The Dutch unemployment figures are taken from various sources of the Central Bureau of Statistics (CBS) and Eurostat's SOCPROT database. They are all expressed in Full-Time Equivalents (FTE). All Dutch employment data, which come from the Central Planning Bureau's *LangeReeksen Boek 1950-1996*, have a similar base. Data on interest rates and the money supply are obtained from the IMF's *International Financial Statistics Yearbook* (from 1990 and 1995 publications). The GDP deflator is also contained herein, and is combined with comparable data from the OECD's STAN database. As a proxy of the short-term interest rate, the money market rate is used (which influences short-term borrowings between financial institutions)[1].

In order to calculate the employment dispersion measure $\hat{\sigma}_t$, we distinguished 20 sectors. These sectors are shown in Table A.1 below.

---

1 As suggested by the IMF itself, cf. the *International Financial Statistics Yearbook 1995* (1995), pp. *xv-xvi.*

**Table A.1**     Sectors used in the calculation of $\hat{\sigma}_t$.

| Number | Description |
|---|---|
| 1 | Agricultural, forestry and fishery products |
| 2 | Food, beverage and tobacco |
| 3 | Textiles and clothing, leather and footwear |
| 4 | Wood, cork and furniture |
| 5 | Paper and printing |
| 6 | Chemical and rubber products, plastics |
| 7 | Metal products, machinery, office and data processing machines, electrical goods, precision and optical instruments |
| 8 | Petroleum and natural gas |
| 9 | Ferrous and non-ferrous ores and metals, minerals and mineral products (excluding petroleum and natural gas) |
| 10 | Production and distribution of electricity, gas, steam and (hot) water |
| 11 | Building and construction |
| 12 | Letting of real estate |
| 13 | Recovery and repair services, wholesale and retail trade services |
| 14 | Maritime and air transport services |
| 15 | Inland and auxiliary transport services |
| 16 | Communication services |
| 17 | Services of credit and insurance companies |
| 18 | Lodging and catering, other market services |
| 19 | Health care, other non-market services (excluding government services) |
| 20 | Government services |

The time series $\hat{\sigma}_t$, that resulted using employment information on the sectors above, subsequently entered the construction of $\hat{\sigma}_t^p$.