Estimation of the monthly unemployment rate for six domains through structural time series modelling with cointegrated trends

Sabine Krieg and Jan van den Brakel

Discussion paper (201018)

Statistics Netherlands

The Hague/Heerlen, 2010

Explanation of symbols

.	= data not available
*	= provisional figure
**	= revised provisional figure
x	= publication prohibited (confidential figure)
–	= nil or less than half of unit concerned
–	= (between two figures) inclusive
0 (0,0)	= less than half of unit concerned
blank	= not applicable
2008–2009	= 2008 to 2009 inclusive
2008/2009	= average of 2008 up to and including 2009
2008/'09	= crop year, financial year, school year etc. beginning in 2008 and ending in 2009
2006/'07–2008/'09	= crop year, financial year, etc. 2006/'07 to 2008/'09 inclusive

Due to rounding, some totals may not correspond with the sum of the separate figures.

Publisher
Statistics Netherlands
Henri Faasdreef 312
2492 JP The Hague

Prepress
Statistics Netherlands - Grafimedia

Cover
TelDesign, Rotterdam

Information
Telephone +31 88 570 70 70
Telefax +31 70 337 59 94
Via contact form: www.cbs.nl/information

Where to order
E-mail: verkoop@cbs.nl
Telefax +31 45 570 62 68

Internet
www.cbs.nl

ISSN: 1572-0314

60083201018 X-10

Estimation of the Monthly Unemployment Rate for Six Domains through Structural Time Series Modelling with Cointegrated Trends

Sabine Krieg and Jan van den Brakel

Summary: The monthly unemployment rate is based on the data of the Dutch Labour Force Survey. In this paper a structural time series model is developed and applied to estimate the monthly unemployment rate for six domains. The estimation results under this model are compared with the generalized regression estimates and the estimates under some simpler models. With univariate structural time series models, information from other time periods is borrowed to improve the precision of the estimates. Further improvements are possible by borrowing information from other domains in a multivariate structural time series model. It turns out that the trends of the six domains are cointegrated. Only three common trends have to be estimated, which means that the information from other domains is used in an efficient way. Further improvements can be achieved by modelling outliers and by modelling the seasonal component in an efficient way. The standard error of the estimates is approximately halved by the time series approach, compared to the generalized regression estimator.

Keywords: Structural time series models, Small area estimation, Unemployment rate, Cointegration

1. Introduction

The purpose of the Dutch Labour Force Survey (LFS) is to produce reliable information about the situation on the labour market. Until 1999 the LFS was a continuously conducted cross-sectional survey. In October 1999 this survey changed to a rotating panel design. The estimation procedure is based on the generalized regression (GREG) estimator to produce figures about employment and unemployment. This estimator is widely applied by national statistical institutes since it is always approximately design unbiased, see e.g. Chambers et al. (2006). GREG estimators, however, have relatively large design variances in the case of small sample sizes. The sample size of the Dutch LFS is too small to produce reliable information about the monthly employed and unemployed labour force with the GREG estimator. In the past, each month the samples observed in the preceding three months were used to estimate moving averages about the labour market situation. Furthermore, quarterly and annual figures about the labour market situation are published at a more detailed level.

In the situation of small sample sizes, small area estimators can be used to improve the precision of direct estimates. These model based estimators generally have smaller variances than GREG estimators because they explicitly rely on statistical models to borrow sample information that is observed in preceding periods or in other domains. The first approach is called borrowing strength over time, the latter borrowing strength over space. For a comprehensive overview of small area estimation techniques, see Rao (2003).

Since a few years, Statistics Netherlands investigates how small area estimators can be used in official statistics. For the Dutch LFS, two lines of model based small area estimation are explored (Boonstra et al., 2008). One approach is based on mixed models, where information from other domains is borrowed to improve the annual municipal figures of the Dutch LFS. Another approach is the use of a multivariate structural time series model, where information from other time periods and from other domains is borrowed to improve monthly unemployment rates at a high aggregation level. The LFS is conducted continuously in time and the monthly unemployment rate is correlated with the unemployment rate in the preceding periods. Direct estimators, like the GREG estimator, only use the data observed in a particular month to estimate the monthly unemployment rate. The precision of the GREG estimates can be improved by using data observed in preceding periods through explicit modeling of time series obtained by repeatedly conducted surveys. This approach dates back to Scott and Smith (1974), who proposed to consider the true value of the finite population parameter as a realization of a stochastic process that can be described with a time series model. Pfeffermann and Bleuer (1993) and Pfeffermann and Burck (1990) proposed a multivariate structural time series model to borrow strength over time and space. Time series models that explicitly account for rotating panel designs are developed by Pfeffermann (1991). Other key references are Pfeffermann et al. (1998), Pfeffermann and Tiller (2006) and Tiller (1992).

The design of the LFS is a rotating panel design, where the respondents are interviewed face to face in the first wave and they are re-interviewed by phone in four subsequent waves. In Van den Brakel (2005) it is explained how the times series approach can be applied to the LFS, where the model accounts for the rotating panel design of this survey. In Van den Brakel and Krieg (2007, 2008, 2009) a multivariate model that accounts for the rotating panel design of the LFS is applied to estimate monthly unemployment figures at the national level. This multivariate model is implemented in the production process in 2010 where it is applied to estimate monthly figures not only at the national level but also at six domains (men and women in three age classes). In this model, no information in space is borrowed.

In Van den Brakel and Krieg (2006) a model is applied to estimate monthly unemployment rates for six domains, where only the first wave of the survey is used. There, information over space is borrowed by modelling the covariance between components of different domains. In Krieg and van den Brakel (2008) different multivariate models to estimate the monthly unemployment rate for six domains are developed and compared. Different possibilities to model the covariance between the trend components are compared in combination with different models for the seasonal component and outlier modelling. The variance of the GREG estimates is not used in the models.

These results are extended in the present paper. Again, the data obtained in the first wave of the panel is used. The models considered in this paper account for the variances of the GREG estimates. Furthermore, the model allows for different covariances between the trend components of the domains. It turns out that the trends are cointegrated, and that it is sufficient to estimate three common trends. The interpretation of this result is discussed.

The model is developed using the data of the LFS from January 1996 until December 2007. As a form of cross-validation, the model is then applied to the data of 2008.

In Section 2 the LFS is summarized. In Section 3, multivariate structural time series models for the monthly unemployment rate for six domains are described. The estimation results of the models are presented in Section 4. The fact that the trends are cointegrated is discussed in Section 5. The paper concludes with a discussion in Section 6.

2. The Dutch Labour Force Survey

2.1 Sample design

Each month a sample of addresses is selected from which during the data collection households are identified that can be regarded as the ultimate sampling units. The target population of the LFS consists of the non-institutionalised population aged 15 years and over residing in the Netherlands. The sampling frame is a list of all known occupied addresses in the Netherlands, which is derived from the municipal basic registration (in Dutch: gemeentelijke basis administratie). The LFS is based on a stratified two-stage cluster design of addresses. Strata are formed by geographical regions. Municipalities are considered as primary sampling units and addresses as secondary sampling units. In the first stage a sample of municipalities is drawn with first order inclusion probabilities proportional to the number of addresses. At the second stage a sample of minimal 12 addresses is drawn without replacement from each selected municipality. All households residing on an address, up to a maximum of three, are included in the sample (in the Netherlands, there is generally one household per address).

Until September 1999, the LFS was a continuous survey with a gross sample size of about 10000 addresses (except for July and August, where the gross sample sizes were halved). At that time, addresses with persons registered at the employment office were oversampled and addresses with only persons aged 65 years and over were undersampled. In October 1999, the LFS changed from a continuous survey to a rotating panel design. The monthly gross sample size averaged about 8000 addresses commencing that moment and gradually declined to about 6500 addresses in 2008. In July and August 2000 and 2001 the gross sample size were halved. Addresses with persons registered at the employment office are not oversampled in the period 2000 – 2010. Since 2008, addresses with young persons or with migrants are oversampled.

During the period that the LFS was a cross-sectional survey and in the first wave of the panel design, data are collected by means of computer assisted personal interviewing (CAPI). For all members of the selected households, demographic variables are observed. For the target variables only persons aged 15 years and over are interviewed. When a household member cannot be contacted, proxy interviewing is allowed by members of the same household. Households, in which one or more of the selected persons do not respond for themselves or in a proxy interview, are treated as non-responding households. Since the LFS has changed to a rotating panel design, the respondents aged 15 through 64 years are re-interviewed four times at quarterly intervals by means of computer assisted telephone interviewing (CATI). The data from these re-interviews are not used in this paper.

2.2 Estimation procedure

The weighting procedure of the LFS is based on the GREG estimator (Särndal et al., 1992). The inclusion probabilities reflect the sampling design described above as well as the different response rates between geographic regions. The weighting scheme is based on a combination of different socio-demographic categorical variables. The integrated method for weighting persons and families of Lemaître and Dufour (1987) is applied to obtain equal weights for persons belonging to the same household. Finally, a bounding algorithm proposed by Huang and Fuller (1978) is applied to avoid negative weights.

Target parameters about the employed and unemployed labour force are defined as population totals or as ratios of two population totals. The unemployment rate, which is investigated in this paper, is defined as the ratio of the total unemployment and the total labour force. The monthly sample size of the LFS is too small to publish reliable monthly figures using the GREG estimator because the variance is too high. Therefore estimates about the employed and unemployed for the preceding 13 weeks were published each month.

2.3 Monthly estimates based on monthly data

In this research project, estimates for the monthly unemployment rate based on the separate monthly samples are produced with a slightly modified version of the estimation procedure currently used at Statistics Netherlands. According to Eurostat rules, periods of 4 or 5 weeks are used instead of calendar months. The sample size that is available for the separate month, requires that the weighting scheme for the GREG estimator currently used in the production process is reduced, and is given by

Age(5)Gender + *Region(44)* + *Gender* × *Age(21)* + (2.1)
Age(5) × *Marital Status* + *Ethnicity(8).*

Age(5)Gender is a classification in 8 classes which is based on 5 age classes and the second, third and fourth age class is itemized to gender. *Region(44)* is a classification in 44 geographic regions. *Age(i)* is a classification in *i* age classes. *Marital Status* is a classification in 2 classes which distinguish between married and not married. *Ethnicity(8)* is a classification in 8 classes.

The unemployment rate is estimated for the following six domains:

1. Men, 15-24 year,
2. Women, 15-24 year,
3. Men, 25-44 year,
4. Women, 25-44 year,
5. Men, 45-64 year,
6. Women, 45-64 year.

These estimates are defined as

$$Y_{t,d} = \frac{t_{y,t,d}}{t_{z,t,d}},$$

with $t_{y,t,d}$ and $t_{z,t,d}$ the GREG estimates for the unemployed labour force and the labour force at time *t* for domain *d*.

Since $Y_{t,d}$ is the ratio of two GREG estimators, an expression for the estimator of the variance of $Y_{t,d}$ is given by

$$\mathrm{var}(Y_{t,d}) = \frac{1}{(t_{z,t,d})^2} \sum_{h=1}^{H} \frac{n_{h,t}}{n_{h,t}-1} \left(\sum_{k=1}^{n_{h,t}} (w_k e_{k,t,d})^2 - \frac{1}{n_{h,t}} \left(\sum_{k=1}^{n_{h,t}} w_k e_{k,t,d} \right)^2 \right), \quad (2.2)$$

with

$$e_{k,t,d} = \sum_{l=1}^{m_k} \left[(y_{k,l,t,d} - \mathbf{x}_{k,l}^T \mathbf{b}_{y,d}) - Y_{t,d} (z_{k,l,t,d} - \mathbf{x}_{k,l}^T \mathbf{b}_{z,d}) \right].$$

Here $y_{k,l,t,d}$ is a binary variable taking value one if the *l*-th person belonging to the *k*-th household belongs to the unemployed labour force and to domain *d* at time *t* and zero otherwise, $z_{k,l,t,d}$ a binary variable taking value one if the *l*-th person of the *k*-th household belongs to the labour force and to domain *d* at time *t* and zero otherwise, $\mathbf{x}_{k,l}$ a vector with the auxiliary information of the *l*-th person belonging to the *k*-th household used in the weighting scheme of the GREG estimator, $\mathbf{b}_{y,d}$ and $\mathbf{b}_{z,d}$ the regression coefficient of the regression function of $y_{k,l,t,d}$ respectively $z_{k,l,t,d}$ on $\mathbf{x}_{k,l}$, w_k the regression weight of household *k*, $n_{h,t}$ the number of completely responding households of stratum *h=1,…H,* at time *t*, and m_k the number of persons aged 15 years and over belonging to the *k*-th household. Persons belonging to the same household have equal weights due to the application of the integrated method for weighting persons and families of Lemaître and Dufour (1987). Formula (2.2) is the variance estimation procedure implemented in Bascula to approximate the variance of the ratio of two GREG estimators, Nieuwenbroek and Boonstra (2002).

3. A structural times series model for six domains

3.1 Model specification

Let θ_t denote the population parameter at time *t*, e.g. the unemployment rate. Direct estimators, like the Horvitz-Thompson estimator or the GREG estimator, assume that θ_t is a fixed but unknown parameter. Under this design-based approach, an estimator for θ_t for cross-sectional surveys only uses the data observed at time *t*. Scott and Smith (1974) proposed to consider the population parameter θ_t as a realization of a stochastic process that can be described with a time series model. Under this assumption, data observed in preceding periods $t-1$, $t-2$, ..., can be used to improve the estimator for θ_t, even in the case of non-overlapping sample surveys. In the context of small area estimation this is called borrowing strength in time. Sample information from different domains can be used to further improve the domain estimates, which is known as borrowing strength in space. One possible approach is to allow for random area and random time effects in a linear mixed model, and apply a composite estimator like the BLUP or EBLUP, see e.g. Rao and Yu (1994). Pfeffermann and Burck (1990) and

Pfeffermann and Bleuer (1993) proposed to borrow strength in time and space by applying a multivariate structural time series model. In this approach, information in space is borrowed by modelling the correlation between the model parameters of the domains.

Each month a GREG estimate is computed for each of the six domains mentioned in Section 2.3. Let $Y_{t,d}$ denote the GREG estimate for $\theta_{t,d}$, the unemployment rate of domain *d* at time *t*, using the estimation procedure described in Section 2.3. So each month a vector $\mathbf{Y}_t = (Y_{t,1}, Y_{t,2}, Y_{t,3}, Y_{t,4}, Y_{t,5}, Y_{t,6})^T$ is observed. This vector can be modelled as

$$\mathbf{Y}_t = \boldsymbol{\theta}_t + \mathbf{e}_t, \qquad (3.1)$$

with $\boldsymbol{\theta}_t = (\theta_{t,1}, \theta_{t,2}, \theta_{t,3}, \theta_{t,4}, \theta_{t,5}, \theta_{t,6})^T$ and $\mathbf{e}_t = (e_{t,1}, e_{t,2}, e_{t,3}, e_{t,4}, e_{t,5}, e_{t,6})^T$ the vector of the corresponding survey errors for each estimate.

With a structural time series model, the population parameter $\boldsymbol{\theta}_t$ can be decomposed in four components, i.e.:

$$\boldsymbol{\theta}_t = \mathbf{L}_t + \mathbf{S}_t + (\mathbf{x}_{t,1}^T \boldsymbol{\beta}_1, ..., \mathbf{x}_{t,6}^T \boldsymbol{\beta}_6)^T + \boldsymbol{\varepsilon}_t, \qquad (3.2)$$

where $\mathbf{L}_t = (L_{t,1}, L_{t,2}, L_{t,3}, L_{t,4}, L_{t,5}, L_{t,6})^T$ denotes the vector of the trend components for the six domains and $\mathbf{S}_t = (S_{t,1}, S_{t,2}, S_{t,3}, S_{t,4}, S_{t,5}, S_{t,6})^T$ the vector of the seasonal components. With the third component it is possible to use auxiliary information in the model. In this paper, this component is used to include the outliers in the model. Then $\mathbf{x}_{t,d} = (x_{t,1,d}, ..., x_{t,K_d,d})^T$ is a K_d-dimensional vector for every domain, where K_d is the number of outliers in domain *d*. $x_{t,k,d}$ is one if the *k*-th outlier in domain *d* occurs in time period *t* and zero otherwise. $\boldsymbol{\beta}_d = (\beta_{1,d}, ..., \beta_{K_d,d})^T$ is a K_d-dimensional vector with time independent regression coefficients $\beta_{k,d}$. Finally $\boldsymbol{\varepsilon}_t = (\varepsilon_{t,1}, \varepsilon_{t,2}, \varepsilon_{t,3}, \varepsilon_{t,4}, \varepsilon_{t,5}, \varepsilon_{t,6})^T$ is the vector of the irregular components. An alternative way to deal with outliers is to consider them as missing observations which results in a more parsimonious model. Model (3.2) is used to test whether the regression coefficients $\beta_{k,d}$ are significant.

The trend component for each domain is based on a stochastic process that describes how the unemployment rate gradually evolves over time. A widely applied stochastic model for the trend component is the smooth trend model (see for example Durbin and Koopman, 2001), which is defined by the following set of equations:

$$\begin{aligned}
&L_{t,d} = L_{t-1,d} + R_{t-1,d}, \\
&R_{t,d} = R_{t-1,d} + \eta_{R,t,d}, \\
&E(\eta_{R,t,d}) = 0, \\
&Cov(\eta_{R,t,d}, \eta_{R,t',d'}) = \begin{cases} \sigma_{R,d}^2 & \text{if} \quad t = t' \quad \text{and} \quad d = d' \\ 0 & \text{if} \quad t \neq t' \\ \varsigma_{R,d,d'} & \text{if} \quad t = t' \quad \text{and} \quad d \neq d'. \end{cases}
\end{aligned} \qquad (3.3)$$

The parameters $L_{t,d}$ and $R_{t,d}$ are referred to as the trend and the slope parameter respectively. Using this model, information from the past about the long term development is used to improve the estimates. When $\varsigma_{R,d,d'} = 0$, no information from other domains about the long term development is borrowed. With $\varsigma_{R,d,d'} \neq 0$ sample information from other domains is used to construct a trend. This way, the estimates are further improved.

In previous research (Van den Brakel and Krieg, 2009), the local linear trend models is investigated as an alternative. In this model, a disturbance is added to the first equation of (3.3). A significance test showed in (Van den Brakel and Krieg, 2009) that this disturbance can be assumed to be zero. Therefore, only the smooth trend model is used in this paper.

In this paper, the so called trigonometric seasonal model is used to model the seasonal component which is defined as:

$$S_{t,d} = \sum_{j=1}^{6} S_{t,d,j}, \qquad (3.4)$$

$$S_{t,d,j} = \cos\left(\frac{j\pi}{6}\right) S_{t-1,d,j} + \sin\left(\frac{j\pi}{6}\right) S^*_{t-1,d,j} + \omega_{t,d,j},$$

$$S^*_{t,d,j} = -\sin\left(\frac{j\pi}{6}\right) S_{t-1,d,j} + \cos\left(\frac{j\pi}{6}\right) S^*_{t-1,d,j} + \omega^*_{t,d,j},$$

for $j = 1,...,5$ and

$$S_{t,d,6} = -S_{t-1,d,6} + \omega_{t,d,6},$$

$$E(\omega_{t,d,j}) = 0, \; E(\omega^*_{t,d,j}) = 0,$$

$$Cov(\omega_{t,d,j}, \omega_{t',d',j'}) = Cov(\omega^*_{t,d,j}, \omega^*_{t',d',j'})$$
$$= \begin{cases} \sigma^2_{\omega,d,j} & \text{if} \quad t = t' \quad \text{and} \quad j = j' \quad \text{and} \quad d = d' \\ 0 & \text{if} \quad t \neq t' \quad \text{or} \quad j \neq j' \\ \varsigma_{\omega,d,d',j} & \text{if} \quad t = t' \quad \text{and} \quad j = j' \quad \text{and} \quad d \neq d', \end{cases}$$

$$Cov(\omega_{t,d,j}, \omega^*_{t',d',j'}) = 0.$$

Using this model, information from the past about the seasonal pattern is used to improve the estimates. Similarly as for the trend, it is possible to model the covariances between the seasonal components. Then the seasonal effects for the different domains change more or less simultaneously, depending on the estimated covariances. Because the seasonal effects are almost time invariant, not much information can be borrowed from other domains this way. Furthermore, under the trigonometric seasonal model this will result in a rather complex model with a large number of hyperparameters. Therefore, $\varsigma_{\omega,d,d',j} = 0$ is chosen.

With the simplification of $\varsigma_{\omega,d,d',j} = 0$, six hyperparameters have to be estimated for each domain in model (3.4). This makes the model quite complex, but also more flexible. In the literature, it is suggested that the model can be simplified by taking $\sigma^2_{\omega,d,j} = \sigma^2_{\omega,d}$ for all j (see Harvey, 1989). This simplification seems not appropriate for the data used in this paper, as will be seen in Section 5.

In the case of $\varsigma_{R,d,d'} = 0$ and $\varsigma_{\omega,d,d',j} = 0$, six univariate models are estimated.

Inserting equation (3.2) into equation (3.1) gives

$$\begin{aligned}\mathbf{Y}_t &= \mathbf{L}_t + \mathbf{S}_t + (\mathbf{x}_{t,1}^T \boldsymbol{\beta}_1, ..., \mathbf{x}_{t,6}^T \boldsymbol{\beta}_6)^T + \mathbf{e}_t + \boldsymbol{\varepsilon}_t \\ &= \mathbf{L}_t + \mathbf{S}_t + (\mathbf{x}_{t,1}^T \boldsymbol{\beta}_1, ..., \mathbf{x}_{t,6}^T \boldsymbol{\beta}_6)^T + \mathbf{v}_t\end{aligned} . \quad (3.5)$$

In this application, the disturbance $\nu_{t,d}$ is dominated by the survey error of the GREG-estimate $Y_{t,d}$. The variance $Var_{t,d}$ of $Y_{t,d}$ is estimated together with $Y_{t,d}$ itself. So, $\nu_{t,d}$ can be modelled by

$$E(\nu_{t,d}) = 0,$$

$$Cov(\nu_{t,d}, \nu_{t',d'}) = \begin{cases} \sigma_{\nu,d}^2 Var_{t,d} & \text{if } \quad t = t' \quad \text{and} \quad d = d' \\ 0 & \text{if } \quad t \neq t' \quad \text{or} \quad d \neq d'. \end{cases} \quad (3.6)$$

Equation (3.6) allows for non homogeneous sampling variation. It is expected that $\sigma_{\nu,d}^2 \approx 1$. This hyperparameter is added considering that the variation of the survey errors does not follow exactly the estimated variance of the GREG estimator.

The general way to proceed is to put the structural time series model in state-space representation (see Harvey, 1989 and Durbin and Koopman, 2001). A state-space model consists of a measurement equation and a transition equation. The measurement equation, sometimes also called the observation equation, specifies how the observed time series depends on a linear combination of the unobservable state variables, e.g. the trend, seasonal and regression component:

$$\mathbf{Y}_t = \mathbf{Z}_t \boldsymbol{\alpha}_t + \mathbf{v}_t,$$

where $\boldsymbol{\alpha}_t$ is the state vector with unobservable state variables and $\mathbf{Z}_t$ a known design matrix that specifies the linear relationship between the observations and the elements of the state vector. The state vector $\boldsymbol{\alpha}_t$, the design matrix $\mathbf{Z}_t$, and the vector of disturbances $\mathbf{v}_t$ are given by

$$\boldsymbol{\alpha}_t = (\alpha_{t,1} \quad ... \quad \alpha_{t,6}),$$

with $\boldsymbol{\alpha}_{t,d} = (L_{t,d}, R_{t,d}, S_{t,d,1}, S_{t,d,1}^*, ..., S_{t,d,6}, \beta_{t,1,d}, ..., \beta_{t,K_d,d})^T$,

$$\mathbf{Z}_t = blockdiag(\mathbf{Z}_{t,1} \quad ... \quad \mathbf{Z}_{t,6}),$$

with $\mathbf{Z}_{t,d} = (1,0,1,0,1,0,1,0,1,0,1,0,1, x_{t,1,d}, \ldots, x_{t,K_d,d})$,

and

$$E(\mathbf{v}_t) = \mathbf{0},$$

$$Cov(\mathbf{v}_t, \mathbf{v}_{t'}) = \begin{cases} \mathbf{H}_t & \text{if } \quad t = t' \\ \mathbf{O} & \text{if } \quad t \neq t' \end{cases}$$

with $\mathbf{H}_t = diag(\sigma_{\nu,1}^2 Var_{t,1}, ..., \sigma_{\nu,6}^2 Var_{t,6})$. Here **0** and **O** denote a vector and a matrix respectively, with each element zero.

The transition equation, sometimes also called the system equation, specifies how the state vector evolves over time, and is given by

$$\boldsymbol{\alpha}_t = \mathbf{T}\boldsymbol{\alpha}_{t-1} + \boldsymbol{\eta}_t ,$$

with

$$E(\boldsymbol{\eta}_t) = \mathbf{0} ,$$

$$Cov(\boldsymbol{\eta}_t, \boldsymbol{\eta}_{t'}) = \begin{cases} \mathbf{Q} & \text{if} \quad t = t' \\ \mathbf{O} & \text{if} \quad t \neq t'. \end{cases}$$

The matrix $\mathbf{T}$ is defined as

$$\mathbf{T} = blockdiag(\mathbf{T}_1 \quad ... \quad \mathbf{T}_6),$$

with

$$\mathbf{T}_d = blockdiag(\mathbf{T}_d^R \quad \mathbf{T}_d^S \quad \mathbf{T}_d^\beta),$$

$$\mathbf{T}_d^R = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix},$$

$$\mathbf{T}_d^S = blockdiag(\mathbf{T}_d^{S_1} \quad ... \quad \mathbf{T}_d^{S_6}),$$

where

$$\mathbf{T}_d^{S_j} = \begin{pmatrix} \cos(j\pi/6) & \sin(j\pi/6) \\ -\sin(j\pi/6) & \cos(j\pi/6) \end{pmatrix} \text{ for } j=1,\dots,5, \text{ and}$$

$$\mathbf{T}_d^{S_6} = (-1) \text{ for } j=6,$$

$$\mathbf{T}_d^\beta = \mathbf{I}_{K_d} .$$

Furthermore,

$$\boldsymbol{\eta}_t = (\boldsymbol{\eta}_{t,1}, ..., \boldsymbol{\eta}_{t,6}),$$

with

$$\boldsymbol{\eta}_{t,d} = (0, \eta_{R,t,d}, \omega_{t,d,1}, \omega^*_{t,d,1}, ..., \omega_{t,d,6}, \mathbf{0}_{K_d}).$$

Finally, the covariance matrix of the transition equation is defined by

$$\mathbf{Q} = \begin{pmatrix} \mathbf{Q}_{11} & \cdot & \cdot & \mathbf{Q}_{16} \\ \cdot & \cdot & & \cdot \\ \cdot & & \cdot & \cdot \\ \mathbf{Q}_{61} & \cdot & \cdot & \mathbf{Q}_{66} \end{pmatrix},$$

with

$$\mathbf{Q}_{dd} = diag(0 \quad \sigma^2_{R,d} \quad \sigma^2_{\omega,d,1} \quad \sigma^2_{\omega,d,1} \quad \dots \quad \sigma^2_{\omega,d,1} \quad \mathbf{0}_{K_d}).$$

The matrices $\mathbf{Q}_{dd'}$ with $d \neq d'$ are (13+ K_d) × (13+ $K_{d'}$)-matrices with all elements zero except the second diagonal element, which is $\varsigma_{R,d,d'}$.

When a structural time series model is expressed in the state-space form, the Kalman filter can be applied to obtain optimal estimates for the state vector $\boldsymbol{\alpha}_t$. The Kalman filter is a recursive procedure to obtain optimal estimates for the state vector at time *t* based on the data up to and including time period *t*, and are referred to as the filtered estimates. The filtered estimate of time *t* can be improved by taking account of the information made available after time *t*. This procedure is referred to as smoothing. Several smoothing algorithms are available in the literature. In this paper, the Kalman filter estimates for the state variables are smoothed with the fixed interval smoother, which is a broadly applied smoothing algorithm, and these estimates are referred to as the smoothed estimates. Furthermore, it is possible to make predictions with the Kalman filter. In this paper, the one-step predictions are used, which are estimates for the period $t+1$ given the information up to and including period t.

The hyperparameters σ_*^2 and ς_* are estimated with a maximum likelihood procedure, using a numerical optimization procedure. See Harvey (1989) or Durbin and Koopman (2001) for technical details. In this paper, Ssfpack 3.0 (Koopman et al., 1999b, and Koopman et al., 2008) in combination with Ox (Doornik, 1998) is used for the computations.

The variance-covariance matrix $\mathbf{Q}$ has to be positive semi-definite. To enforce this, the matrix can be written in the form $\mathbf{Q} = \mathbf{ADA}^T$ with $\mathbf{D}$ a diagonal matrix and $\mathbf{A}$ a lower triangular matrix with ones on the diagonal. When $\mathbf{A}$ is written in the form

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & . & . & \mathbf{A}_{16} \\ . & . & & . \\ . & & . & . \\ \mathbf{A}_{61} & . & . & \mathbf{A}_{66} \end{pmatrix},$$

then $\mathbf{A}_{dd}$ are unit matrices of size (13+ K_d) × (13+ K_d) and $\mathbf{A}_{dd'}$ with $d > d'$ are (13+ K_d) × (13+ $K_{d'}$)-matrices with all elements zero except the element in the second row and second column. This element has to be estimated in the maximum likelihood procedure.

3.2 Preliminary results and model improvements

In a preliminary analysis a model without outliers is estimated. The model evaluation (for the used evaluation measures see appendix A.1) shows that the correlogram of the standardized prediction errors of the first domain has a cyclical pattern with a large value at lag 1. Furthermore, the standardized prediction errors of the third and the fifth domain are not normally distributed. For the sixth domain, the standardized prediction errors are heteroscedastic. This is due to two outliers in the first domain (9-2003 and 10-2003), one outlier in the third domain (3-1998), two outliers in the fifth domain (06-2000 and 12-2006) and two outliers in the sixth domain (7-2006 and 3-2007). It can be expected that the model better fits the data when dummy indicator variables for these outliers are included in the time series model.

For the fourth domain, the standardized prediction error of 2-1999 is notably large. This value does not cause suspicious evaluation results, nevertheless it is possibly an improvement of the model to treat this observation as an outlier.

The (smoothed) estimates of the parameters $\beta_{k,d}$ are significantly different from zero for all outliers.

A further preliminary analysis shows that some of the harmonics of the seasonal components are small and not significantly different from zero. These harmonics are shown in Table 3.1.

Table 3.1: Seasonal harmonics which are not significantly different from zero

Domain	*Seasonal harmonics*
1	$S_{t,1,4}, S^*_{t,1,4}, S_{t,1,5}, S^*_{t,1,5}$
2	$S_{t,2,1}, S^*_{t,2,1}, S_{t,2,5}, S^*_{t,2,5}, S_{t,2,6}$
3	$S_{t,3,2}, S^*_{t,3,2}, S_{t,3,3}, S^*_{t,3,3}, S_{t,3,4}, S^*_{t,3,4}, S_{t,3,5}, S^*_{t,3,5}$
4	$S_{t,4,3}, S^*_{t,4,3}, S_{t,4,4}, S^*_{t,4,4}, S_{t,4,5}, S^*_{t,4,5}, S_{t,4,6}$
5	$S_{t,5,2}, S^*_{t,5,2}$
6	$S_{t,6,2}, S^*_{t,6,2}, S_{t,6,3}, S^*_{t,6,3}, S_{t,6,4}, S^*_{t,6,4}, S_{t,6,5}, S^*_{t,6,5}, S_{t,6,6}$

This test for significance is executed in two steps. In the first step, it is tested whether the harmonic is constant over time. Subsequently, it is tested whether the harmonic is significantly different from zero for at least one month.

In cases where two or more harmonics are not significantly different from zero, it is also tested whether the sum of the harmonics is significantly different from zero. After this test, it is decided that the seasonal harmonics in Table 3.2 can be removed from the model.

Table 3.2: Seasonal harmonics which are removed from the model

Domain	*Seasonal harmonics*
1	$S_{t,1,4}, S^*_{t,1,4}, S_{t,1,5}, S^*_{t,1,5}$
2	$S_{t,2,5}, S^*_{t,2,5}, S_{t,2,6}$
3	$S_{t,3,2}, S^*_{t,3,2}, S_{t,3,3}, S^*_{t,3,3}, S_{t,3,5}, S^*_{t,3,5}$
4	$S_{t,4,3}, S^*_{t,4,3}, S_{t,4,4}, S^*_{t,4,4}, S_{t,4,5}, S^*_{t,4,5}, S_{t,4,6}$
5	$S_{t,5,2}, S^*_{t,5,2}$
6	$S_{t,6,2}, S^*_{t,6,2}, S_{t,6,4}, S^*_{t,6,4}$

The model is adapted by removing the corresponding components from equation (3.5) and from $\boldsymbol{\alpha}_t$, $\mathbf{Z}_{t,d}$, $\mathbf{T}_d$, $\boldsymbol{\eta}_{t,d}$, and $\mathbf{Q}$ in the state-space representation.

3.3 Comparison of seven models

In order to investigate the influence of different model definitions, seven models are compared in this paper. The influence of outlier modelling and of the simplifications of the seasonals is investigated by comparing the following three models:

1. a model without outlier modelling, and with all seasonal harmonicss included,
2. a model with outlier modelling, and with all seasonal harmonics included,
3. a model with outlier modelling and with the seasonal harmonics of Table 3.2 removed.

In these models, the correlation between the disturbances of slope parameters is modelled as described in (3.3) where for different covariances $\varsigma_{R,d,d'} \neq 0$ is allowed. In order to investigate the influence of these correlations, the models are compared with 4 parsimonious models:

4. a model with covariances $\varsigma_{R,d,d'} = 0$, without outlier modelling, and with all seasonal harmonics included,
5. a model with covariances $\varsigma_{R,d,d'} = 0$, with outlier modelling, and with seasonal harmonics of Table 3.2 removed,
6. a model assuming that the correlations between the disturbances of the slopes of all domains are equal, by taking $\varsigma_{R,d,d'} \Big/ \sqrt{\sigma^2_{R,d}\sigma^2_{R,d'}} = \vartheta_R$ for all $d \neq d'$. No outlier modelling is included, all seasonal harmonics are included,
7. a model assuming that the correlations between the disturbances of the slopes of all domains are equal, by taking $\varsigma_{R,d,d'} \Big/ \sqrt{\sigma^2_{R,d}\sigma^2_{R,d'}} = \vartheta_R$ for all $d \neq d'$. Outlier modelling is included, the seasonal harmonics of Table 3.2 are removed.

The following table gives an overview of the models:

Table 3.3: Overview over the models

Model	*Covariance between disturbances of slopes*	*Outliers modelling*	*Some seasonal harmonics removed*
1	Different	No	No
2	Different	Yes	No
3	Different	Yes	Yes
4	Zero	No	No
5	Zero	Yes	Yes
6	Equal correlations	No	No
7	Equal correlations	Yes	Yes

The estimation results of these seven models are presented in the next section.

4. Estimation results

In this section, estimation results for the seven models described in Section 3 are presented. First, the filtered estimates of the different models are graphically compared with each other and with the GREG estimates, where the data up to and including 2007 are used. Then the standard errors are computed in Section 4.2. In Section 4.3, model evaluation results are presented. Finally, the models are applied to the data of 2008 in Section 4.4.

4.1 Filtered estimates

The model estimates obtained under the time series model are compared with the GREG estimates. If the model were used to produce an estimate for month t for regular publication purposes, the estimates of both the model parameters and the estimates of the hyperparameters would be based on the information available at that moment. Therefore the GREG estimates are compared with the filtered estimates for the unemployment rate. In an ideal comparison the hyperparameters are also based on the information available at that moment. In this application, changes of the estimates of the hyperparameters appeared to be small if new observations become available. Since this hardly affects the filtered estimates for the unemployment rate, it was decided to base the hyperparameters on the entire series.

Outlier detection will generally not be applied to the final observation of the series and will therefore not directly affect the filtered estimates for the unemployment rate as they will be published in a regular production process. This technique is, nevertheless, worthwhile to consider since detection and modelling of outliers in the past will improve the model and therefore indirectly result in better filtered estimates for the unemployment rate.

The results of one or two domains are presented and discussed in the text; the results of the other domains are included in appendix A.3.

The model estimates obtained under the time series model are compared with the GREG estimates. In these comparisons, the filtered estimates are used because they are based on the complete set of information that would be available if the model were used to produce an estimate for month t for regular publication purposes. The results of one or two domains are presented and discussed in the text; the results of the other domains are included in appendix A.3.

In Figure 4.1, the GREG estimates are compared with the filtered model estimates under model 1 for domain 1.

Figure 4.1: GREG estimates and filtered estimates under model 1 for domain 1

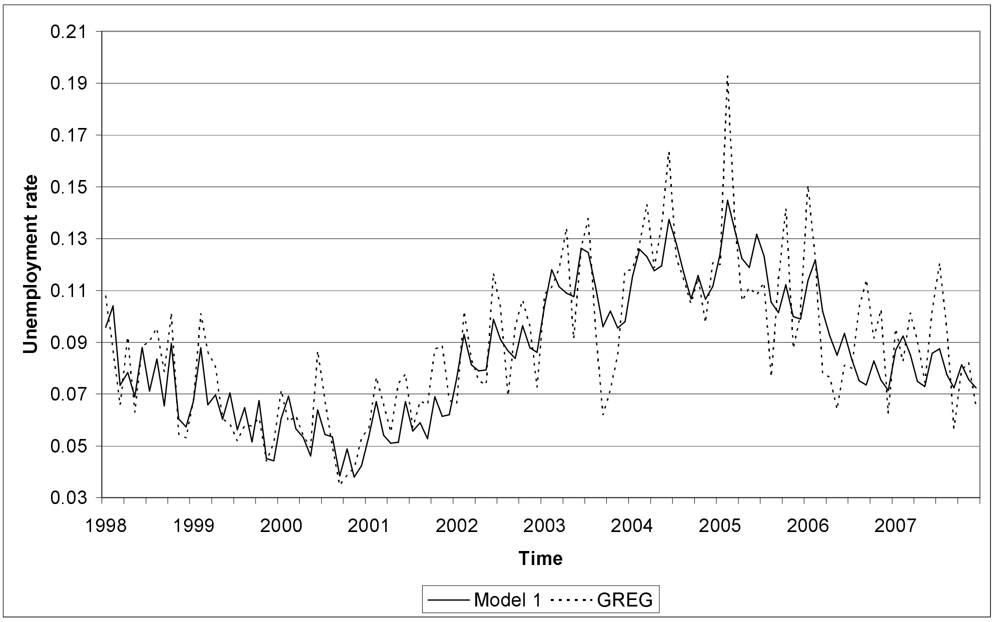
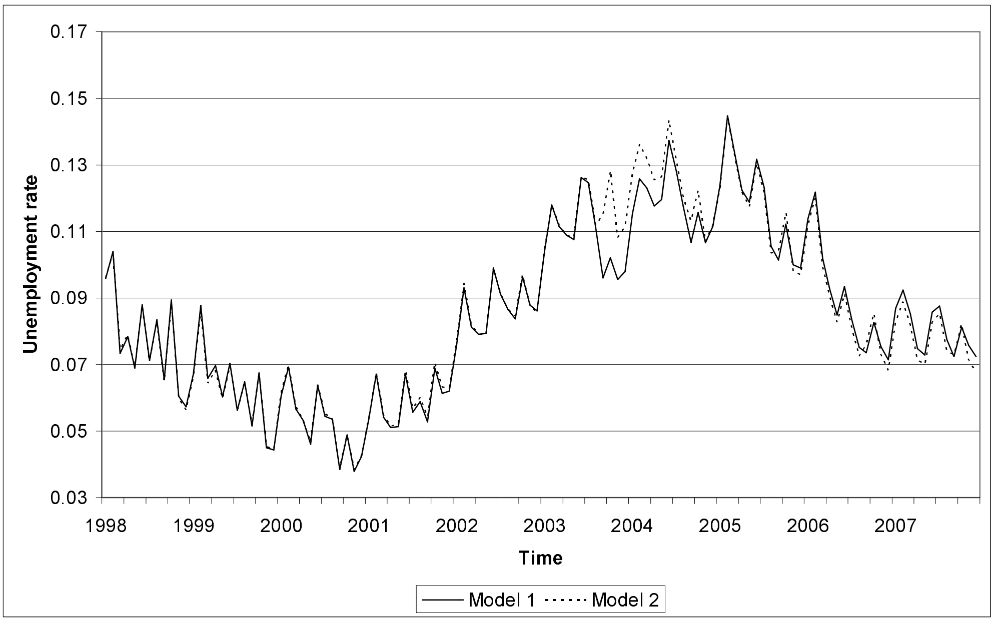


Figure 4.1 shows that the filtered estimates are at the same level as the GREG estimates. They partly follow the fluctuations in the GREG series, since these fluctuations are considered as seasonal effects under the model. Nevertheless a substantial part of the irregularities in the series of the GREG estimates are flattened out, since they are considered as survey errors under the time series model. It appears that the filtered estimates are not seriously biased.

In Figure 4.2 the filtered estimates of model 1 and model 2 are compared. This figure shows the influence of the outlier modelling. For the time periods preceding the periods where the outliers are modelled, the filtered estimates under both models are equal. For the time periods where the outliers are modelled and for some months after these time periods, the filtered estimates under model 2 are larger than the estimates under model 1, because model 2 is not influenced by the outliers. Some minor differences in the years after the periods where the outliers are modelled are caused by slightly different estimates of the hyperparameters.

The Figures A.6 – A.10 in appendix A.3 show that the filtered estimates under model 2 are smaller than the estimates under model 1 for some other domains. In these domains, the outliers are extremely large, whereas in domain 1, the outliers are extremely small. The comparison of Figure 4.2 with the figures in the appendix shows furthermore that the differences between the estimates under model 1 and model 2 for domain 1 are larger than for the other domains. This has to do with the fact that in domain 1, two outliers are found in two successive months.

Figure 4.2: Filtered estimates under model 1 and model 2 for domain 1



In Figure 4.3 the filtered estimates of model 2 and model 3 are compared for domain 1. This figure shows the influence of a more parsimonious model for the seasonal effects. Only small and not significant seasonal components are removed under model 3. Therefore, the model estimates under model 2 and model 3 are very similar. In domain 4, more seasonal components are removed under model 3 resulting in smoother estimates (see Figure 4.4).

Figure 4.3: Filtered estimates under model 2 and model 3 for domain 1

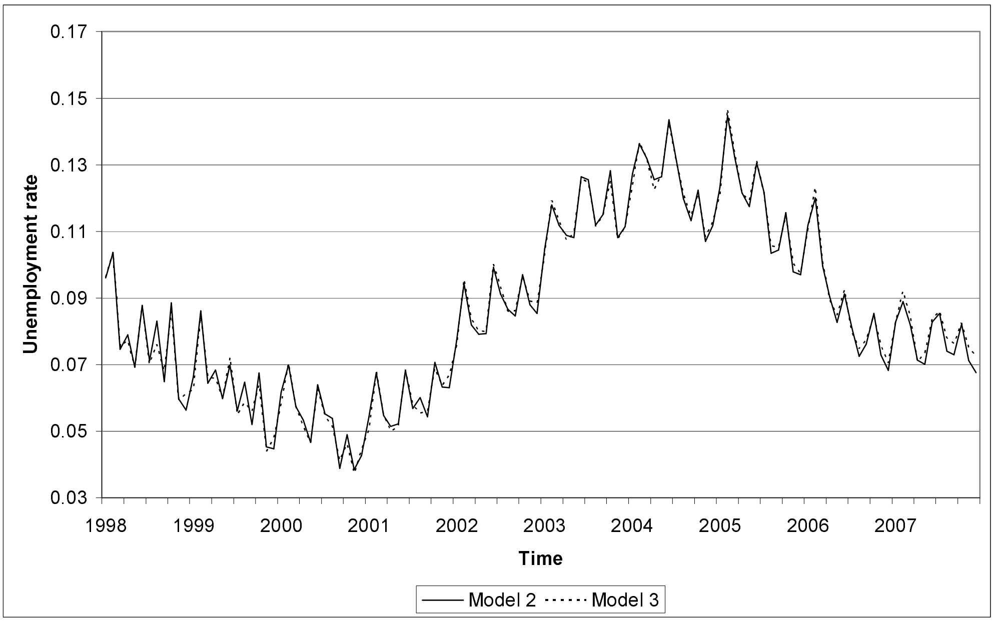
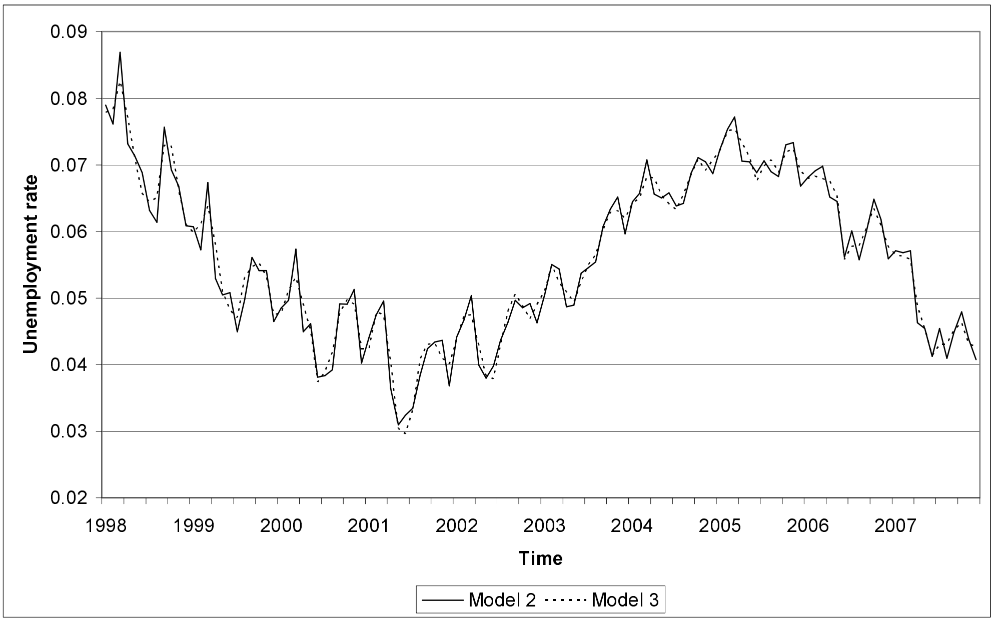
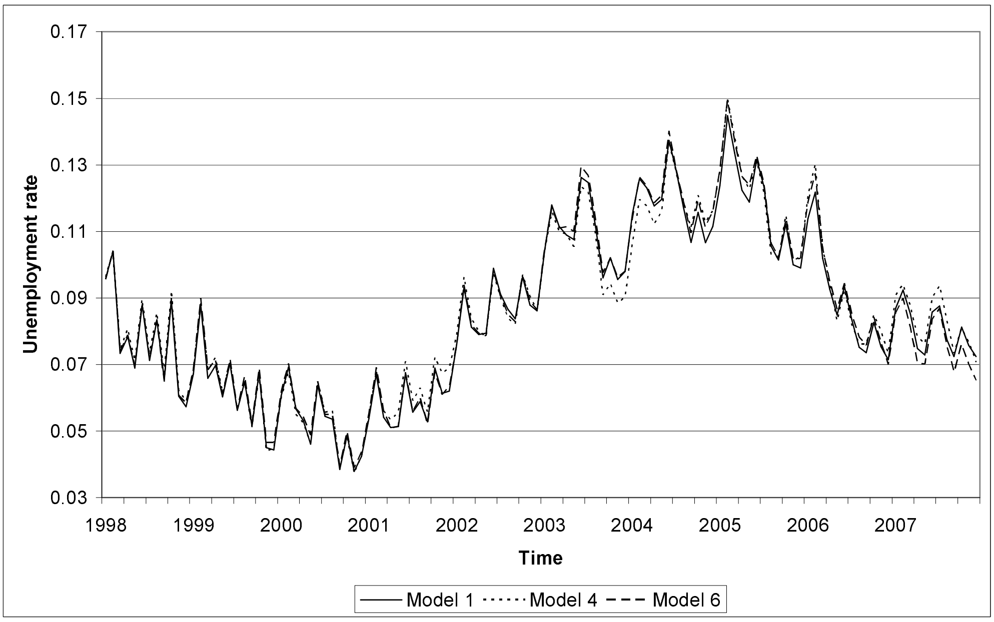


Figure 4.4: Filtered estimates under model 2 and model 3 for domain 4



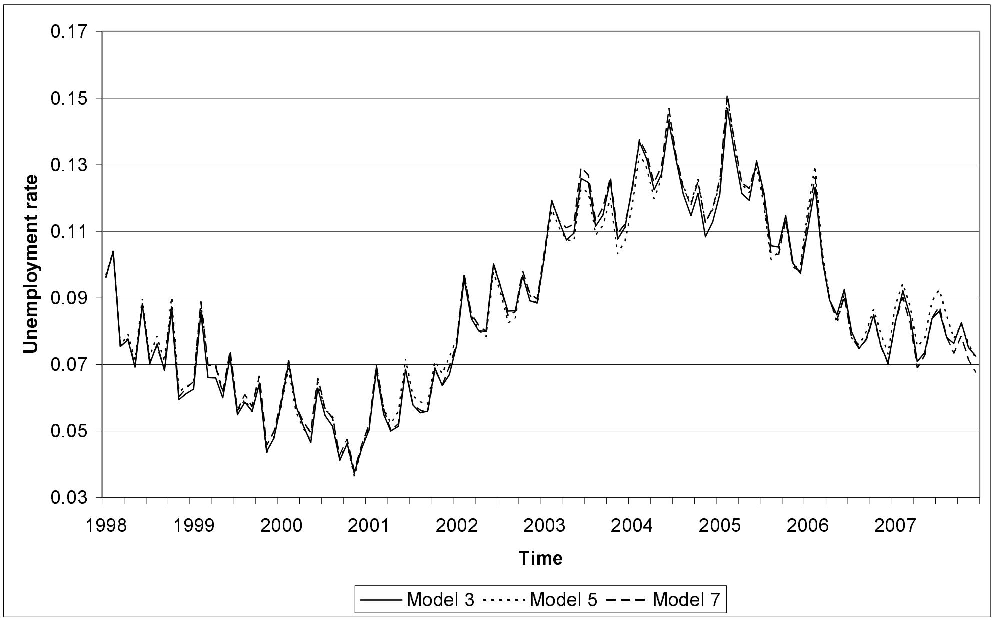
In Figure 4.5 the filtered estimates under model 1, model 4, and model 6 are compared for domain 1. This figure shows the influence of different ways of borrowing information from other domains. Under model 1 and model 6, the estimates of the trend are influenced by the other domains due to the correlations between the disturbances of the slope parameters, whereas under model 4, the estimates of the trend are based on the information from domain 1 only. Because the estimates of the correlations under model 1 are not the same as under model 6, the estimates of the trend differ under these models. As expected, the seasonal patterns under the three models are similar.

Figure 4.5: Filtered estimates under model 1, model 4, and model 6 for domain 1



Similarly, Figure 4.6 also shows the influence of different ways of borrowing information from other domains, where the filtered estimates under model 3, model 5, and model 7 are compared for domain 1.

Figure 4.6: Filtered estimates under model 3, model 5, and model 7 for domain 1



Generally, it can be concluded that the filtered model estimates are at the same level as the GREG estimates, and therefore, they are not seriously biased. Based on the figures, it cannot be decided which model fits the data best and has, therefore, to be preferred. In the next subsections, the model estimates are further compared.

4.2 Standard errors

In Table 4.1, the mean of the standard errors over 2006-2007 of the GREG estimates for all domains is presented in the second column. In the other columns, the ratio of the mean standard errors over 2006-2007 of the filtered model estimates to the mean standard error of the GREG estimates is given. It can be concluded that the standard error of the GREG estimates is approximately halved by the time series approach. Comparing model 1 with model 4 and model 6, or model 3 with model 5 and model 7, it can be concluded that the precision can be further improved by borrowing information from other domains, and that the estimates are more precise when the model allows for different correlations.

The precision of the estimates is improved by modelling outliers (compare model 1 with model 2). For domain 2, no outliers are modelled, therefore, no improvement of the precision is expected there. The standard error is even larger under model 2 for domain 2 under the influence of outlier modelling in the other domains. Comparing model 2 with model 3 shows that the precision is improved with a more parsimonious model with less seasonal components.

Table 4.1: Mean standard errors over 2006-2007 of GREG estimates and ratio of mean standard errors of filtered model estimates and mean standard errors GREG

	Mean standard error GREG	Ratio mean standard error and column 2 Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Domain 1	0.0188	0.41	0.40	0.38	0.45	0.42	0.43	0.40
Domain 2	0.0215	0.47	0.49	0.45	0.56	0.55	0.52	0.51
Domain 3	0.0056	0.48	0.44	0.42	0.56	0.51	0.53	0.49
Domain 4	0.0075	0.48	0.43	0.40	0.54	0.49	0.52	0.47
Domain 5	0.0067	0.41	0.41	0.41	0.50	0.49	0.47	0.46
Domain 6	0.0095	0.49	0.45	0.42	0.54	0.47	0.51	0.44

The precision of the estimates is one aspect in the model selection process, however, if the model is missspecified, the estimates can be biased and then the accuracy is worse. The figures in Section 4.1 already show that the model estimates are not seriously biased. In the model evaluation in Section 4.3 this is further investigated.

4.3 Model diagnostics

In appendix A.1, the assumptions underlying the state-space model are described together with some standard methods for model evaluation to test for these assumptions. As described in Section 3.2, some model assumptions are violated under model 1, which is due to some outliers. As expected, most model assumptions are satisfied under the models including outliers (models 2, 3, 5, and 7). Under model 4 and 6, the same model assumptions are violated as under model 1.

For domain 1, the absolute values of at least two sample autocorrelation functions are larger than the bound of 0.179 of the 95% confidence interval under all models. Whereas under the models without outlier modelling the sample autocorrelation function of lag 1 is extremely large (around 0.3) and there is a clear cyclical pattern in the correlogram, under the models with outlier modelling the sample autocorrelation function of lag 1 is less extreme (around 0.2). Also the sample autocorrelation functions of larger lags are not extremely large. Therefore it can be concluded that also for domain 1, the outlier modelling improves the model evaluation results.

By modelling a third outlier in the first domain, the sample autocorrelation functions become smaller, but the model estimates seem to be biased: the filtered estimated are smaller than the GREG-estimates over a longer period. Therefore the third outlier is not included in the model.

4.4 Cross-validation

Cross-validation is a widely applied measure of the predictive power of a model. Details are given in Appendix A2. A large value of the mean of the prediction errors (MPE, formula A.2) indicates that the model estimates are biased. In the second column of Table 4.2, the standard error of the mean of the 24 GREG estimates of the period 2006-2007 is given. In the other columns, the ratio of the MPE to this standard error is computed.

Table 4.2: Ratios of MPE under 7 models and the standard error of the mean of the GREG estimates, period 2006-2007

	Standard error mean GREG estimates	Ratio MPE and column 2 Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Domain 1	0.0039	-2.54	-2.77	-2.57	-0.96	-0.86	-2.35	-2.43
Domain 2	0.0045	2.82	2.83	2.86	3.08	3.05	2.65	2.71
Domain 3	0.0012	-0.65	-0.17	-0.21	-0.46	-0.57	-1.24	-1.08
Domain 4	0.0015	0.02	-0.21	-0.25	0.58	0.13	0.44	0.15
Domain 5	0.0014	-0.75	0.25	0.22	-0.13	1.06	-0.66	0.46
Domain 6	0.0020	-0.78	-1.22	-1.10	0.84	0.81	-1.01	-0.84

Because of the sampling errors of the GREG estimates and ignoring the fact that the predictions are estimates with some uncertainty, a ratio between -1.96 and 1.96 would be expected. Because of this uncertainty a slightly larger or smaller ratio can be expected. In most cases, the absolute value of this ratio is smaller than 1.96. The absolute values of the ratios in Table 4.2 are larger than 1.96 for domain 2 under all models and for domain 1 under the multivariate models. Figure 4.7 shows the predictions under model 1, model 4, and model 6 compared with the GREG estimates for domain 1. The figure shows that the predictions are larger than the GREG estimates for several months, for example February until June 2006, and smaller during other periods of months, for example August until November 2006. During most months of 2007, the predictions under model 4 are closer to the GREG estimates than the predictions under model 1 and 6. Therefore, the MPE under model 4 is closer to zero. Under the multivariate models 1 and 6, the level of the estimates of domain 1 are influenced by the estimates of the other domains due to the correlation between the disturbances of the slope parameters.

Figure 4.8 shows the predictions under model 1, model 4, and model 6 compared with the GREG estimates for domain 2. There, the predictions are larger than the GREG estimates during most months of the considered years.

Figure 4.9 and Figure 4.10 show that the filtered estimates under model 1 are closer to the GREG estimates than the predictions.

Taking into account that the MPE's are not extremely large and that both the model predictions and the GREG estimates are estimated with some uncertainty, it cannot be concluded from the predictions that the model estimates are biased.

Figure 4.7: GREG estimates and predictions under model 1, model 4, and model 6 for domain 1

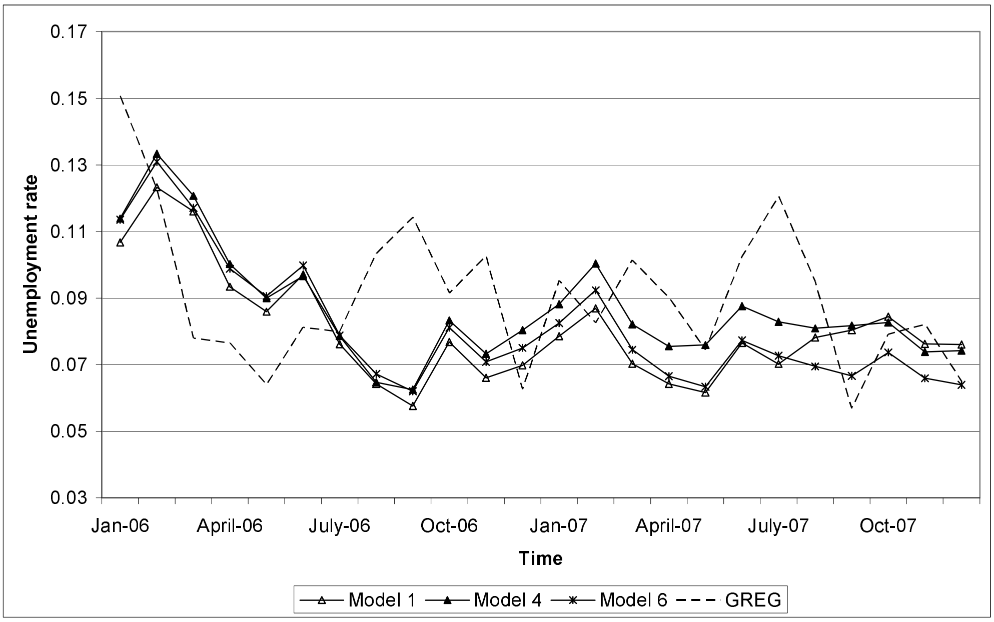


Figure 4.8: GREG estimates and predictions under model 1, model 4, and model 6 for domain 2

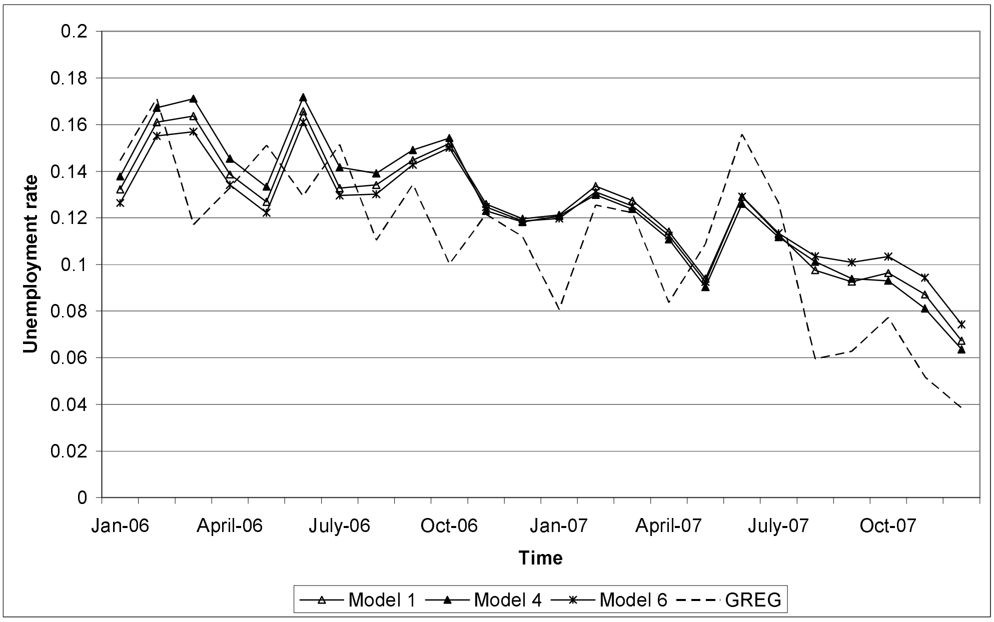


Figure 4.9: GREG estimates and predictions and filtered estimates under model 1, for domain 1

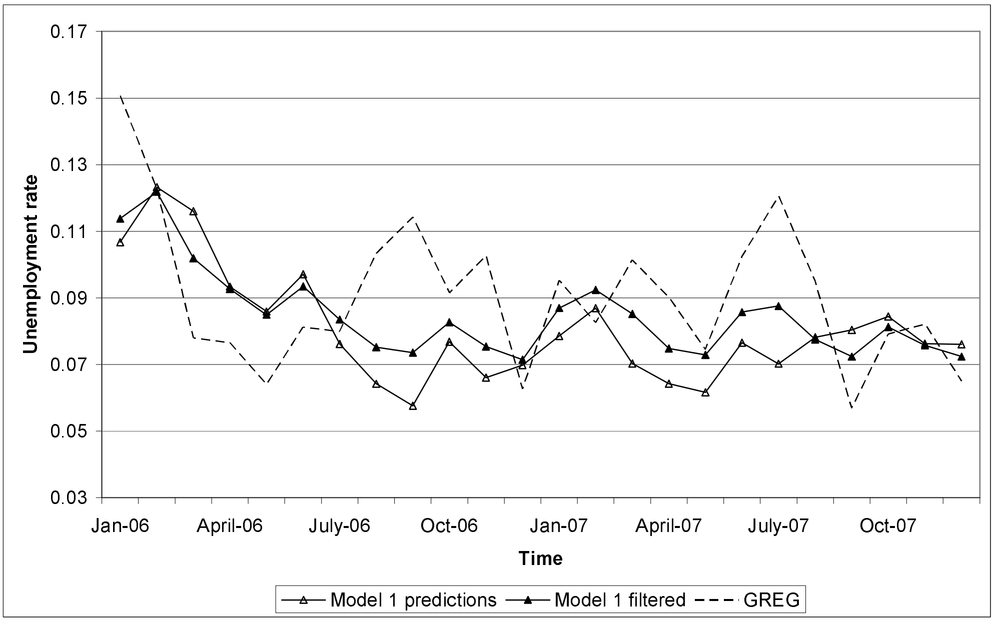
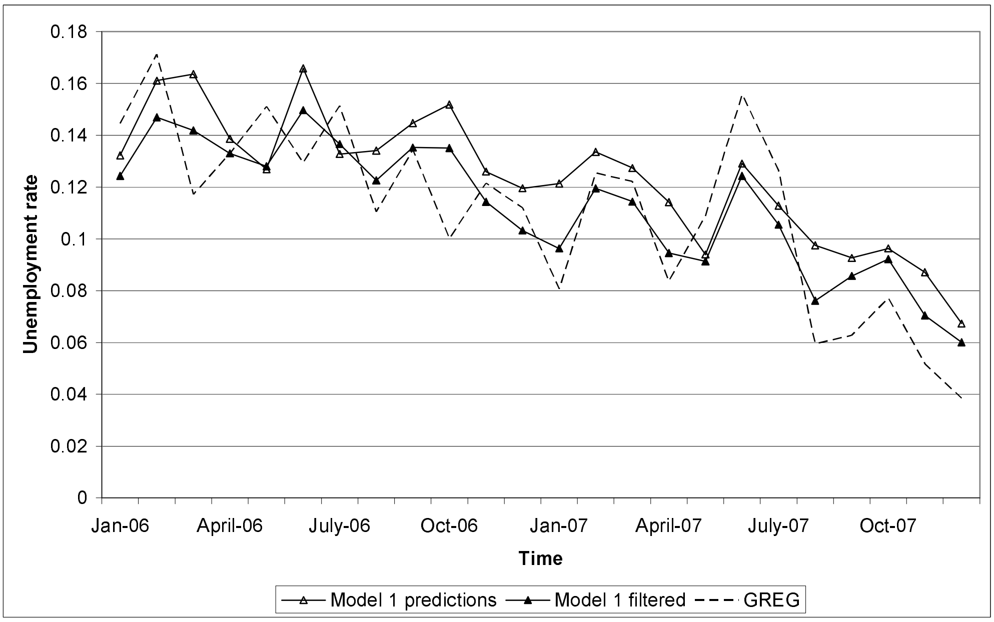


Figure 4.10: GREG estimates and predictions and filtered estimates under model 1, for domain 2



The square root of the mean squared prediction error (RMSPE, formula A.4) under model 1 is shown in the second column of Table 4.3 (period 2006-2007). In the other columns, the ratio of the RMSPE under the other models and the RMSPE under model 1 is computed. It can be concluded that the predictive power is similar for all models. Because the GREG estimates are measured with uncertainty, a slightly smaller RMSPE does not mean that the model is better.

In this application, the results for the mean absolute prediction error (formula A.3) are similar. Therefore, the table is omitted here.

Table 4.3: RMSPE under model 1 and ratio of RMSPE under model 2 – 7 and under model 1, 2006-2007

	RMSPE	Ratio of RMSPE and column 2					
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Domain 1	0.0267	1.01	0.96	0.90	0.94	0.95	0.99
Domain 2	0.0264	0.98	0.96	1.03	1.00	1.04	1.00
Domain 3	0.0075	0.97	0.96	1.01	0.97	1.04	0.98
Domain 4	0.0090	1.03	1.00	1.08	1.06	1.02	1.02
Domain 5	0.0092	0.98	1.01	1.01	1.02	1.00	1.00
Domain 6	0.0171	0.95	0.93	1.01	0.91	0.98	0.92

4.5 Cross validation using the data of 2008

As a form of cross-validation, the models are applied on the data of 2008. This way it can be checked whether the results change if new data are added.

Comparison of the filtered estimates with the GREG estimates and with each other shows similar results as for the period 1996-2007. The model estimates are at the same level as the GREG estimates, whereas most irregularities in the series of the GREG estimates are flattened out since they are considered as survey errors under the models. The differences between the models are small.

Also, the results of the standard tests hardly change. Again, some of the model assumptions are not met under model 1, 4, and 6, caused by outliers. Under model 3, most model assumptions are met. The only exception is the first domain, where some of the sample autocorrelation functions are slightly larger than the computed bound. This is similar as in the case where the data of 2008 are excluded.

As expected, the improvement of the precision is similar in this period as in the period of 2006-2007, compare Table 4.4 and Table 4.1.

Table 4.4: Mean standard errors over 2008 of GREG estimates and relative mean standard errors of filtered model estimates

		Ratio mean standard error and column 2						
	Mean standard error GREG	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Domain 1	0.0180	0.41	0.39	0.38	0.45	0.41	0.43	0.40
Domain 2	0.0186	0.50	0.51	0.47	0.58	0.56	0.53	0.52
Domain 3	0.0047	0.51	0.46	0.43	0.58	0.53	0.55	0.50
Domain 4	0.0062	0.52	0.47	0.44	0.58	0.52	0.56	0.50
Domain 5	0.0057	0.45	0.44	0.44	0.54	0.52	0.51	0.49
Domain 6	0.0076	0.55	0.49	0.46	0.59	0.50	0.57	0.48

Table 4.5 shows the MPE for 2008. Similarly as in Table 4.2, the standard error of the mean of the 12 GREG estimates of the period 2008 is given in the second column. In the other columns, the ratio of the MPE and this standard error is computed. Whereas this ratio is relatively large for domain 1 under most models for the period 2006-2007, the ratio is small for the period 2008. For domain 2, the ratio is relatively large and positive for the period 2006-2007, whereas it is large and negative for the period 2008. Figure 4.11 shows that the predictions for this domain under model 1, model 2, and model 3 are

slightly smaller than the GREG estimates for most of the months of 2008. The somehow large values of the MPE in the period 2006-2007 seem to be coincidences when the results over 2008 are added. For domain 3, the ratio is relatively large and negative for the period 2008. Figure 4.12 shows that the predictions are smaller than the GREG estimates for around the second half of 2008. Similarly as for the period 2006-2007 (Table 4.2) there is no indication that the model estimates are biased.

Table 4.5: Ratios of MPE under 7 models and the standard error of the mean of the GREG estimates, period 2008

		Ratio MPE and column 2						
	Standard error mean GREG estimates	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Domain 1	0.0053	0.08	-0.17	-0.01	-0.74	-0.70	-1.42	-1.01
Domain 2	0.0054	-2.89	-3.53	-3.30	-3.66	-3.61	-2.51	-2.97
Domain 3	0.0014	-2.67	-2.35	-2.49	-2.11	-2.03	-2.63	-2.53
Domain 4	0.0018	-0.56	-1.04	-1.04	-0.86	-0.89	-0.47	-0.83
Domain 5	0.0017	-0.09	-0.10	-0.13	-0.49	-0.88	0.25	-0.02
Domain 6	0.0022	1.45	1.61	1.46	1.03	1.04	2.40	2.19

Figure 4.11: Predictions under model 1, model 2, and model 3 for domain 2, 2008

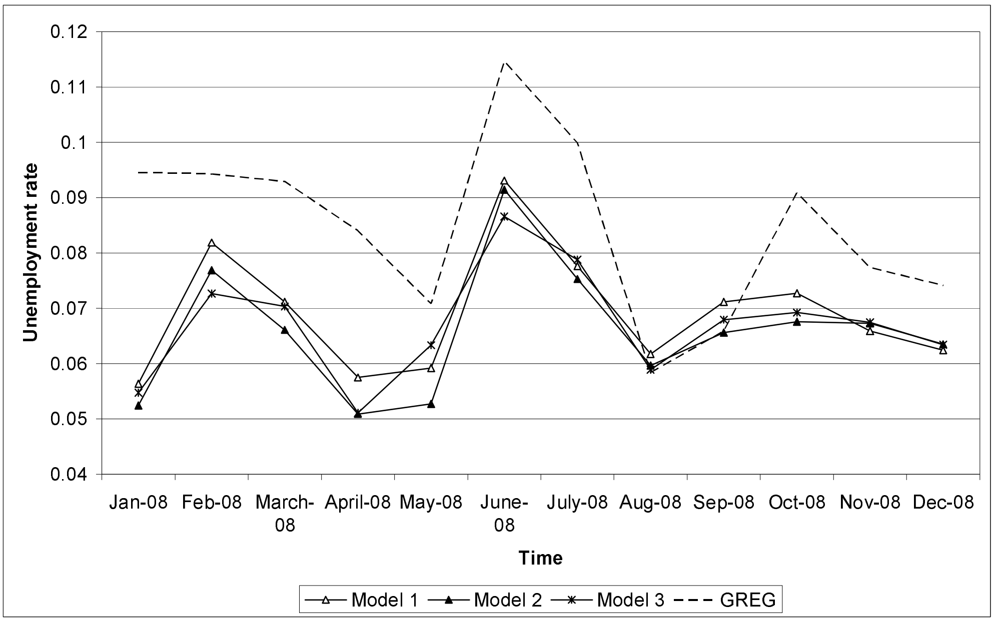


Figure 4.12: Predictions under model 1, model 2, and model 3 for domain 3, 2008

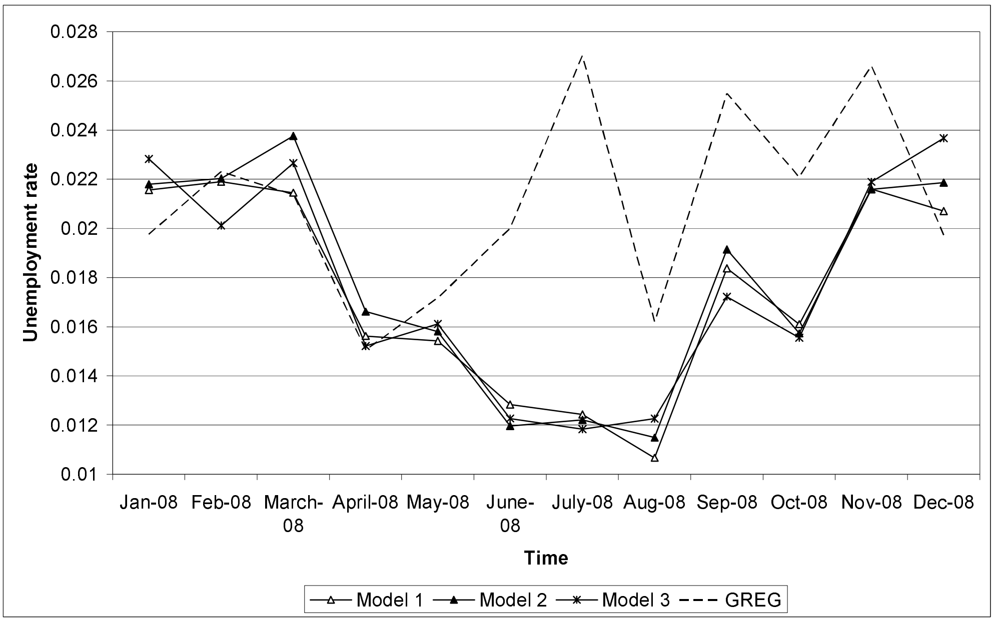


Table 4.6 shows the RMSPE for 2008. Similarly as in Table 4.3, the RMSPE (formula A.4) under model 1 is shown in the second column. In the other columns, the ratio of the RMSPE under the other models and the RMSPE under model 1 is computed. It can be concluded that the predictive power is again similar under all models. Because now a shorter period is considered, the results are more influenced by coincidences.

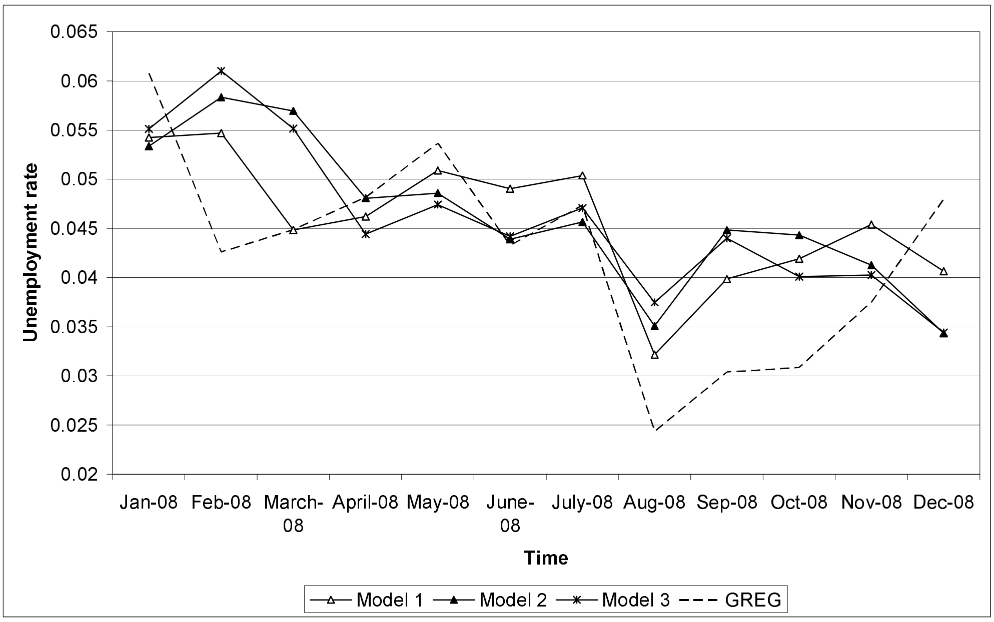
Figure 4.11 shows that for some months (for example February, April and June) the prediction under model 1 is closer to the GREG estimate than the predictions under model 2 and model 3, for domain 2. Therefore, the RMSPE under model 1 is slightly smaller than the RMSPE under model 2 and model 3. Similarly, for domain 6 the predictions under model 1 are closer to the GREG than the predictions under model 2 and model 3 (February, March, August, and September, see Figure 4.13). Therefore, the RMSPE under model 1 is slightly smaller than the RMSPE under model 2 and model 3.

It can be concluded that the predictive power is similar under all models.

Table 4.6: RMSPE under model 1 and ratio of RMSPE under model 2 – 7 and under model 1, 2008

	RMSPE	Ratio of RMSPE and column 2					
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Domain 1	0.0164	1.04	0.98	1.02	0.99	1.08	1.01
Domain 2	0.0194	1.16	1.12	1.23	1.22	0.90	1.01
Domain 3	0.0059	1.02	1.07	0.96	1.04	0.98	1.06
Domain 4	0.0065	1.07	1.00	1.01	1.00	0.99	0.99
Domain 5	0.0076	1.08	1.14	1.01	1.16	1.00	1.12
Domain 6	0.0073	1.37	1.36	0.93	1.33	1.17	1.42

Figure 4.13: Predictions under model 1, model 2, and model 3 for domain 6, for 2008



4.6 Model selection

It is found that for model 3 most model assumptions are met, whether the data from 2008 are included or not. Furthermore, model 3 has the most precise estimates, whereas the predictive power is similar under all models. Therefore, model 3 is selected.

5. Cointegrated trends and other estimation results of the state vector

In this section the estimation results of model 3 are discussed in more detail. In Section 5.1 and 5.2, the data of 1996 -2008 are used. In Section 5.3, it is shown how the results change if a shorter time series of GREG estimates is used as input of the time series model.

5.1 The trend components

It turns out that in matrix $\mathbf{Q} = \mathbf{ADA}^T$ the 31st, the 38th, and the 51st element of the diagonal matrix $\mathbf{D}$ are estimated as zero. These elements belong to the disturbances of the slopes of the forth, fifth and sixth domain. This means that the trends of the six domains are cointegrated and that three common trends are sufficient to describe the unemployment rate of six domains. Model 3 can also be written in the following way:

$$\mathbf{Y}_t = \mathbf{\Phi L}_t^+ + \mathbf{L}_{\Phi,t} + \mathbf{S}_t + (\mathbf{x}_{t,1}^T \boldsymbol{\beta}_1, ..., \mathbf{x}_{t,6}^T \boldsymbol{\beta}_6)^T + \mathbf{v}_t$$

with $\mathbf{L}_t^+ = (L_{t,1}^+, L_{t,2}^+, L_{t,3}^+)$ the vector with three common trends, which are modelled as smooth trends without correlation, $\mathbf{\Phi}$ a 6×3-matrix of standardized factor loadings with ones on the diagonal, i.e. $\varphi_{ii} = 1, i = 1,2,3$, and $\varphi_{ij} = 0$ for $j > i$. The other elements of $\mathbf{\Phi}$ are the non-zero non-diagonal elements of the matrix $\mathbf{A}$, i.e. $\varphi_{2,1} = a_{13,2}$, $\varphi_{3,1} = a_{23,2}, \ldots, \varphi_{6,3} = a_{51,23}$. $\mathbf{L}_{\Phi,t}$ is a vector of six elements, where the first three elements are zero and the last three elements have the form $L_d + R_d t$. The components $\mathbf{S}_t$, $(\mathbf{x}_{t,1}^T \boldsymbol{\beta}_1, ..., \mathbf{x}_{t,6}^T \boldsymbol{\beta}_6)$ and $\mathbf{v}_t$ are defined as in Section 3 (for general information about cointegration under structural time series models see Koopman et. al. 1999a).

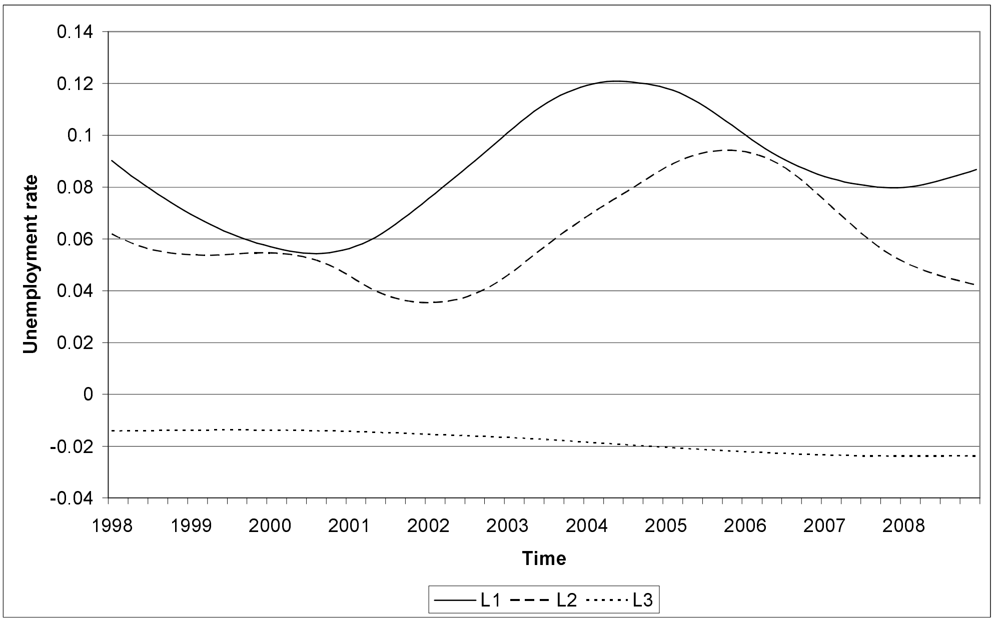
Because of the common trends, information from other domains is borrowed in an efficient way. Shortly spoken, only three independent trends are necessary to describe the trends of the six domains.

The following matrix $\boldsymbol{\Phi}$ is found

$$\begin{pmatrix} 1 & 0 & 0 \\ 0.390 & 1 & 0 \\ 0.397 & 0.246 & 1 \\ 0.302 & 0.299 & -4.128 \\ 0.115 & 0.267 & -1.367 \\ 0.030 & 0.148 & -7.978 \end{pmatrix}.$$

$\mathbf{L}_t^+ = (L_{t,1}^+, L_{t,2}^+, L_{t,3}^+)$ and $\mathbf{L}_{\Phi,t}$ can be computed using $\mathbf{L}_t = \boldsymbol{\Phi}\mathbf{L}_t^+ + \mathbf{L}_{\Phi,t}$. The smoothed estimates of the three common trends $\mathbf{L}_t^+ = (L_{t,1}^+, L_{t,2}^+, L_{t,3}^+)$ are shown in Figure 5.1.

Figure 5.1: Common trends



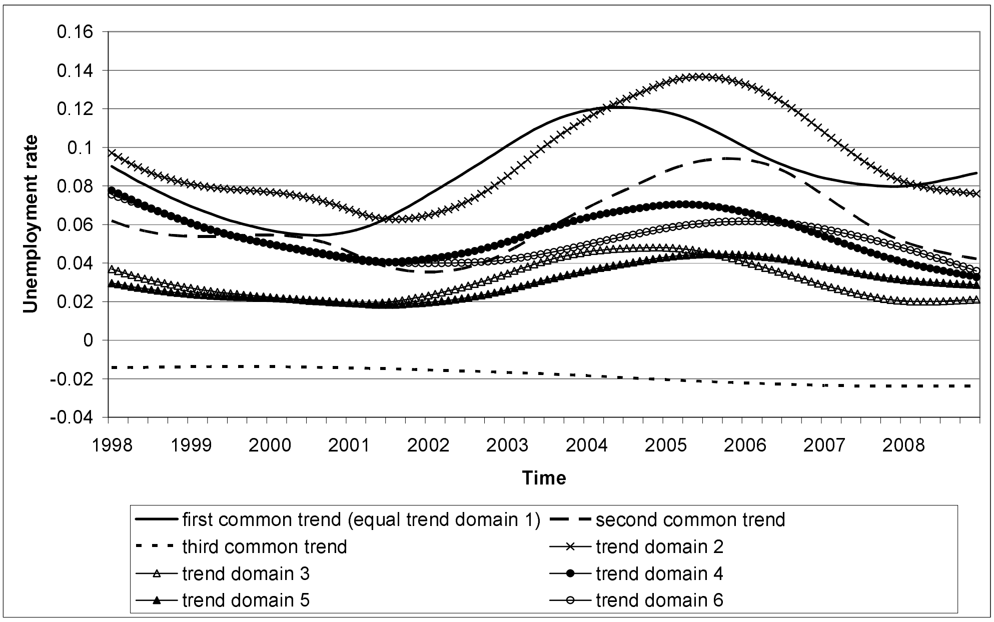
The values for L_d and R_d are given in Table 5.1.

Table 5.1: The values for L_d and R_d

	d		
	4	5	6
R_d	-0.000593	-0.000064	-0.000865
L_d	-0.011427	-0.014903	-0.027200

The fact that the trends are cointegrated offers interesting possibilities of interpretation. In this application, the trend of the 6 domains can (up to the linear correction $L_d + R_d t$) be written as a linear combination of three common trends. The first and the second trend are quite similar, except that the minimum and maximum of the second trend are around 18 months later than the minimum and maximum of the first trend. The third trend is smaller and negative, it is, however, multiplied with relatively large negative constants for the last three domains. The matrix $\mathbf{\Phi}$ shows which weights are given to these three trends to compute the trends of the six domains. Figure 5.2 shows the smoothed estimates of the trends $(L_{t,1}, L_{t,2}, L_{t,3}, L_{t,4}, L_{t,5}, L_{t,6})$ of model 3 together with the smoothed estimates of the common trends $\mathbf{L}_t^+ = (L_{t,1}^+, L_{t,2}^+, L_{t,3}^+)$.

Figure 5.2: Smoothed estimates of trends and common trends



For domain 2, the trend $L_{t,2}$ is a linear combination of the common trends $L_{t,}^+$ and $L_{t,2}^+$, therefore, the maximum of $L_{t,2}$ is between the maxima of $L_{t,}^+$ and $L_{t,2}^+$; because the weight of the second trend $L_{t,2}^+$ is larger, the maximum is close to the maximum of $L_{t,2}^+$. For domain 3 and 4, the weights of $L_{t,}^+$ and $L_{t,2}^+$ are similar, therefore, the maxima of $L_{t,3}$ and $L_{t,4}$ are around halfway between the maxima of $L_{t,}^+$ and $L_{t,2}^+$. For domain 5 and 6, the weight of the second trend $L_{t,2}^+$ is large, therefore, the maxima of $L_{t,5}$ and $L_{t,6}$ are close to the maximum of $L_{t,2}^+$, similarly as for domain 2. The third trend $L_{t,3}^+$ has no clear maximum, therefore, the influence of this trend on the maxima is small. For domain 6 however, with a large weight of $L_{t,3}^+$, the maximum of the trend is later than the maximum of $L_{t,2}^+$. It can be concluded from the matrix $\mathbf{\Phi}$ and from Figure 5.2 that in the considered period, domain 3 and domain 4 (man and women 25-44) develop more or less simultaneously. The same is true for domain 2 (women 15-44), domain 5 and 6 (men and women 45-64). These latest three groups are the last to take advantage from improvements of the labour market.

The matrix $\mathbf{\Phi}$ is not unique. By transforming this matrix, other interpretations are possible (see Koopman et. al. 1999a). Because the main goal of the paper is the estimation of the unemployment rate, this is not worked out.

From the estimation results, the correlations between the disturbances of the slopes can be computed. They are presented in Table 5.2.

Table 5.2 Correlations under the model with different correlations

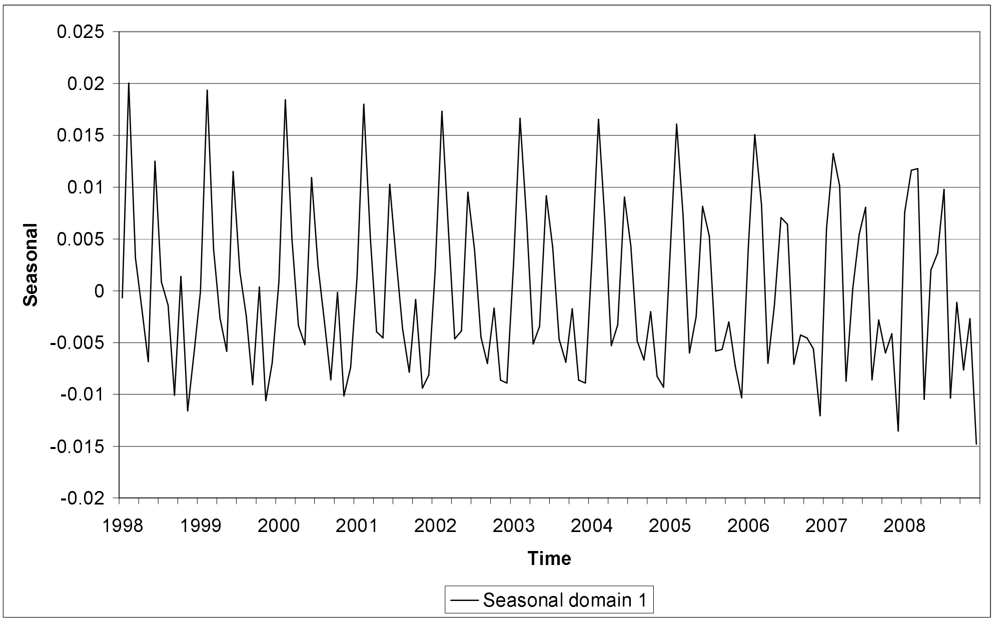
	Domain 2	Domain 3	Domain 4	Domain 5	Domain 6
Domain 1	0.32	0.81	0.61	0.34	0.08
Domain 2		0.81	0.86	0.98	0.44
Domain 3			0.86	0.79	0.23
Domain 4				0.92	0.69
Domain 5					0.60

The table shows that some of the correlations are close to 1, whereas other correlations are small. In the model with equal correlations (model 7), the correlation is estimated as 0.65.

5.2 The seasonal components

The smoothed estimates of the seasonal effect of domain 1 is shown in Figure 5.3.

Figure 5.3: Smoothed estimates of seasonal effect of domain 1 under model 3



The figure shows that the seasonal effect changes gradually over time. The hyperparameter of the sixth harmonic of the seasonal is estimated as 4.47×10^{-7}, which means that this harmonic is allowed to change gradually over time. The hyperparameters of the other harmonics, however, are estimated as zero, which means that the other harmonics are constant over time. The sum of these constant harmonics (first, second and third) is shown in Figure 5.4, the sixth harmonic is shown in Figure 5.5. The sum of the estimates in Figure 5.4 and Figure 5.5 equals the estimates in Figure 5.3. Note that the sixth harmonic is positive in the even months and negative in the uneven months in the first years. In the last years this is reversed. This explains the differences between the estimates of 1998 and 2008 in Figure 5.3.

Figure 5.4: Smoothed estimates of the sum of first, second and third seasonal harmonic of domain 1 under model 3

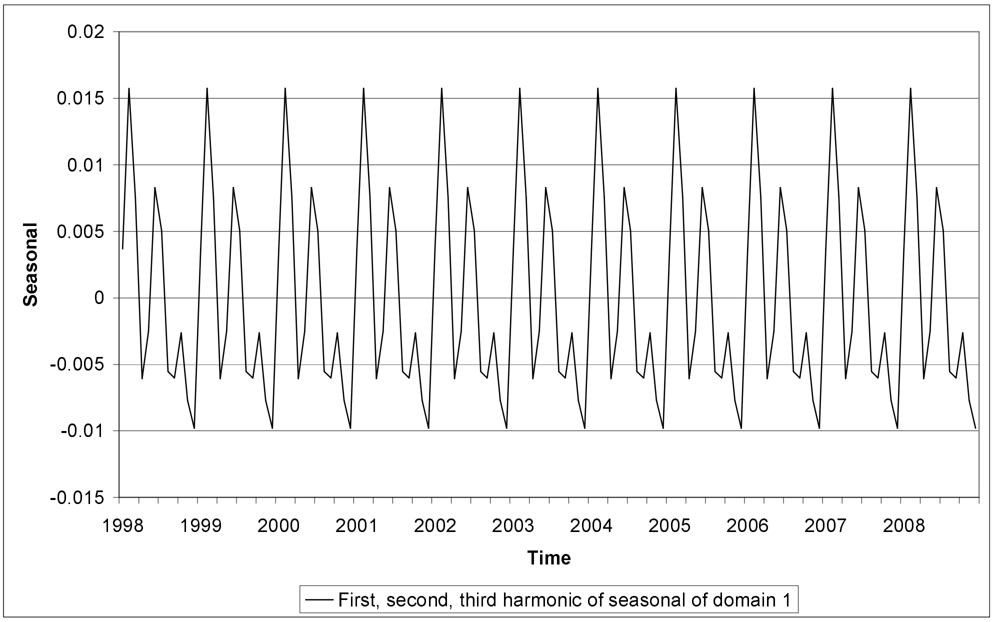
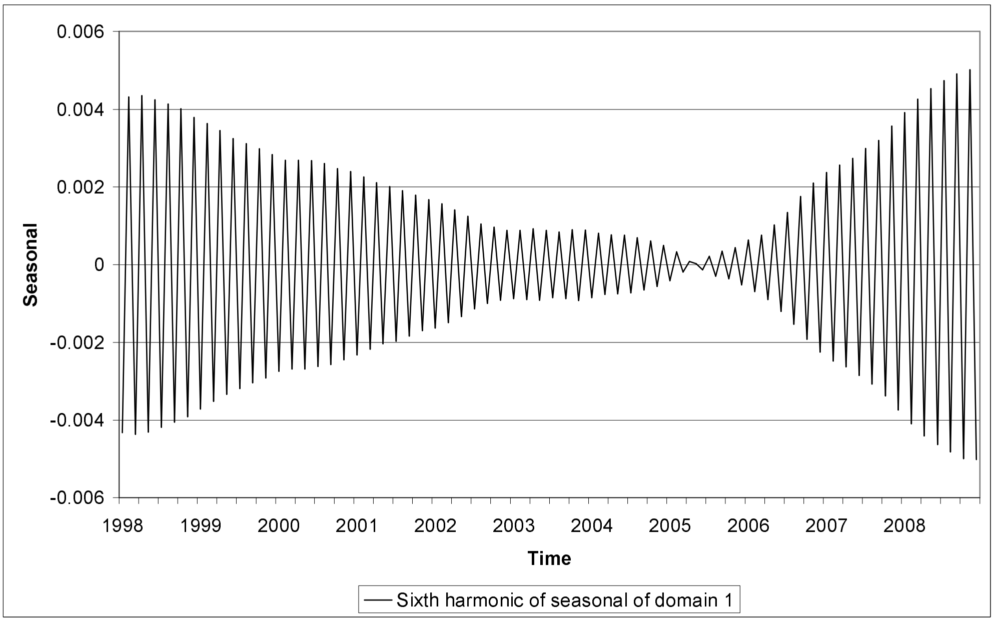


Figure 5.5: Smoothed estimates of sixth seasonal component of domain 1 under model 3



Because some of the seasonal components are allowed to change gradually over time, whereas others are constant, the simplification $\sigma^2_{\omega,d,j} = \sigma^2_{\omega,d}$ in formula (3.4) is not appropriate in this application. This simplification would mean that either all components are constant, or no component is constant.

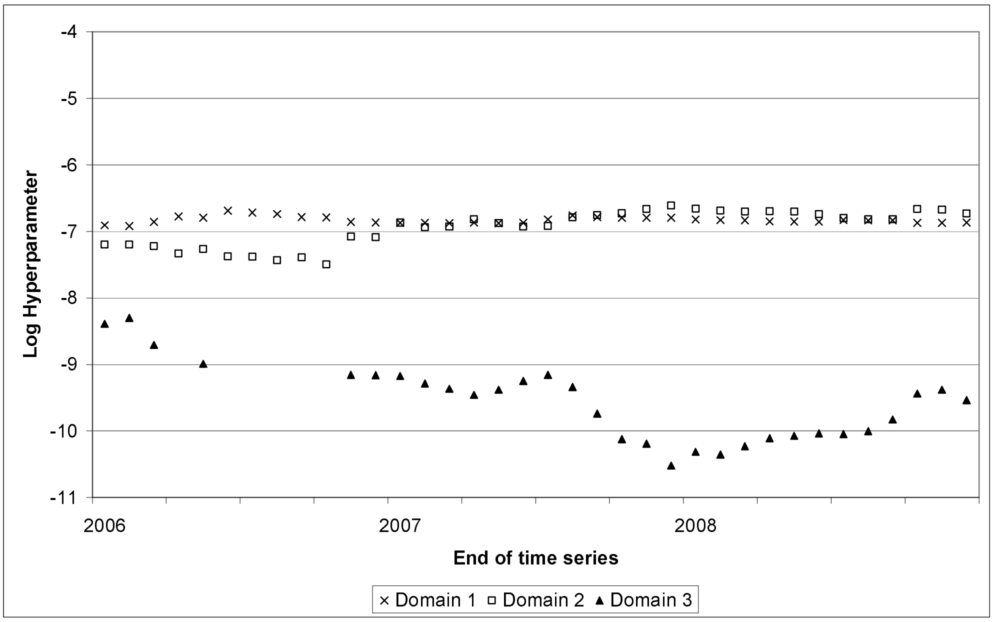
The seasonal effects of most of the other domains also change gradually over time, the figures are added in appendix A.3 (A25-A29).

5.3 Results depend on length of the series

The results shown in Section 5.1 and 5.2 are based on the data 1996 – 2008. In this case, three common trends are necessary. With a shorter or longer series, this can change, since the estimates of the hyperparameter depend on the length of the available series. Figure 5.6 shows the logarithms of the second, the 13th, and the 23rd element of the diagonal matrix $\mathbf{D}$ when the time series from 1996 up to one of the months in 2006-2008 is used as input. These elements belong to the disturbances of the slopes of the first three domains. The 23rd element is zero when the time series ends in April 2006 or June – October 2006.

The 31st, the 38th, and the 51st element of the diagonal matrix $\mathbf{D}$ are estimated as zero also for the shorter series.

Figure 5.6: Logarithms of hyperparameters versus length of time series



The figure shows that the estimates of the hyperparameters are rather stable for the first and second domain. For the third domain, the estimates of the hyperparameter are clearly smaller than for the first two domains. Furthermore, it is less stable. Most importantly, the estimate is zero for some shorter series. This means that in these cases, only two common trends are necessary. Since 2007, three common trends are always necessary. It decreased in the second half of 2007 and increased in 2008.

The hyperparameters $\sigma^2_{v,d}$ of the signal equation (see equation 3.6) are given in Table 5.2. These hyperparameters are indeed estimated close to 1.

Table 5.2: Hyperparameters of the signal equation

Domain	Hyperparameter
1	0.68
2	1.12
3	1.02
4	0.84
5	0.84
6	0.95

6. Discussion and conclusion

In this paper a multivariate structural time series model to estimate the unemployment rate of six domains simultaneously is developed.

With structural time series models, the precision of the estimates can be improved substantially compared with the GREG estimates. This approach is based on the idea that the population parameter of this month is strongly related to the population parameter of the month before. A time series model can be used to describe the development of this population parameter and making advantage from sample information observed in preceding periods to improve the estimates of the actual month. Further, but substantially smaller, improvements are possible by borrowing information from other domains by modelling the correlation between the parameters of the time series models for the domains.

Seven models are compared. The model which is finally selected allows for different correlations between the disturbances of the slope parameters of all domains. Because the slope parameter describes the changes in the trends of the domains, information about the trend of the other domains is borrowed to improve the estimates by these correlations. Outlier modelling is included in the model and the non-significant components of the seasonal effect are removed. The filtered estimates under competing models are quite similar. The model selection is therefore based on the model evaluation and the precision of the model estimates. The selected model is more precise than the competing models. Furthermore, the model evaluation results for this model are satisfying. In this application, the gain in precision by borrowing information from other domains is large, because the trends are cointegrated. Effectively, only three trends have to be estimated in the model. The precision of the model is also improved because some of the harmonics of the seasonal component are removed. The selected model meets the underlying model assumptions for almost all domains because of outlier modelling. Without this outlier modelling, some model assumptions are violated which is demonstrated with some of the competing models.

As a form of cross-validation, the mean of the prediction errors is used to check for biased model estimates, which are not found. It is furthermore expected that the mean of the prediction errors, the mean absolute prediction error and the mean squared prediction error are useful to compare the predictive power of different models. For the models considered in this paper, however, the predictive power is similar. Therefore, these measures are not helpful to distinguish between the considered models.

The model described in this paper does not use the information from the CATI-waves of the rotating panel design of the LFS and it is therefore not appropriate to use this model in the production process to estimate monthly figures about the unemployment rate. A structural time series model that takes advantages of the rotating panel design to estimate the unemployment rate of the total population is described in Van den Brakel and Krieg (2007, 2008, 2009). This model is implemented in the production process in 2010. It is used to estimate figures about the total population and about domains. However, no information is borrowed from other domains in this model. The results from this paper show that by extending this model for six domains the precision of the estimates can be further improved. Further improvement would be possible by using auxiliary information, e.g. information about the registered unemployment.

References

Boonstra, H.J., J.A. van den Brakel, B. Buelens, S. Krieg and M. Smeets (2008). *Towards Small Area Estimation at Statistics Netherlands*. Metron, 56, pp. 21-50.

Brakel, J.A. van den (2005). Small Area Estimators for the Dutch Labour Force Survey using Structural Time Series Models. CBS rapport, bpa nr.: TMO-R&D-2005-05-02-JBRL. Centraal Bureau voor de Statistiek, Heerlen.

Brakel, J.A. van den and S. Krieg (2006). Kleindomeinschatters bij de Enquête beroepsbevolking via tijdreeksmodellen. CBS rapport (in Dutch). bpa nr TMO-R&D-2006-01-20-JBRL. Centraal Bureau voor de Statistiek, Heerlen.

Brakel, J.A. van den and S. Krieg (2007). Modelling Rotation Group Bias and Survey Errors in the Dutch Labour Force Survey. *Proceedings of the section on Survey Research Methods*, American Statistical Association.

Brakel, J.A. van den and S. Krieg (2008). Estimation of the Monthly Unemployment rate through Structural Time Series Modelling in a Rotating Panel Design. Discussion paper 08003, Statistics Netherlands, Heerlen.

Brakel, J.A. van den, and S. Krieg (2009). Estimation of the Monthly Unemployment Rate through Structural Time Series Modelling in Rotating Panel Design. *Survey Methodology* 16, pp. 177-190.

Chambers, R., J.A. van den Brakel, D. Hedlin, R. Lehtonen, and Li-Chun Zhang (2006). Future Challenges of Small Area Estimation, *Statistics in Transition*, 7, pp. 759-769.

Doornik, J.A. (1998). *Object-Oriented Matrix Programming using Ox 2.0*. London: Timberlake Consultants Press.

Durbin, J. and S.J. Koopman (2001). *Time series analysis by state space methods*. Oxford: Oxford University Press.

Harvey, A.C. (1989). *Forecasting, Structural Time Series Models and the Kalman Filter.* Cambridge: Cambridge University Press.

Huang, E.T. and W.A. Fuller (1978). Nonnegative Regression Estimation for Survey Data, *Proceedings of the Social Statistics Session*, American Statistical Association 1978, pp. 300-303.

Koopman, S.J., A.C. Harvey, J.A. Doornik, and N. Shephard (1999a). STAMP: Structural Time Series Analyser, Modeller and Predictor. London: Timberlake Consultants Press.

Koopman, S.J., N. Shephard and J.A. Doornik (1999b). Statistical Algorithms for Models in State Space using SsfPack 2.2. *Econometrics Journal*, 2, pp. 113-166.

Koopman, S.J., N. Shephard and J.A. Doornik (2008). Statistical Algorithms for Models in State Space using SsfPack 3.0. London: Timberlake Consultants Press.

Krieg, S. and J.A. van den Brakel (2008). Estimation of the Monthly Unemployment Rate for Six Domains through Structural Time Series Modelling. CBS rapport, bpa nr SKRG-2008-06-13-DMH. Centraal Bureau voor de Statistiek, Heerlen.

Lemaître, G. and J. Dufour (1987). An Integrated Method for Weighting Persons and Families. *Survey Methodology*, 13, pp. 199-207.

Nieuwenbroek, N. and H.J. Boonstra (2002). Bascula 4.0 Reference Manual. CBS rapport, bpa nr.: 279-02-TMO. Centraal Bureau voor de Statistiek, Heerlen.

Pfeffermann, D. (1991). Estimation and Seasonal Adjustment of Population Means Using Data from Repeated Surveys. *Journal of Business & Economic Statistics*, 9, pp. 163-175.

Pfeffermann, D. and S.R. Bleuer (1993). Robust Joint Modelling of Labour Force Series of Small Areas. *Survey Methodology*, 19, pp. 149-163.

Pfeffermann, D. and L. Burck (1990). Robust Small Area Estimation Combining Time Series and Cross-Sectional Data. *Survey Methodology*, 16, pp. 217-237.

Pfeffermann, D., M. Feder, and D. Signorelli (1998). Estimation of Autocorrelations of Survey Errors with Application to Trend Estimation in Small Areas. *Journal of Business & Economic Statistics,* 16, pp. 339-348.

Pfeffermann, D. and R. Tiller (2006). Small Area Estimation with State Space Models Subject to Benchmark Constraints. *Journal of the American Statistical Association*, 101, pp. 1387-1397.

Rao, J.N.K. (2003). *Small Area Estimation*. New York: Wiley en Sons.

Särndal, C-E., B. Swensson, and J. Wretman (1992). *Model Assisted Survey Sampling.* New York: Springer Verlag.

Scott, A.J. and T.M.F. Smith (1974). Analysis of Repeated Surveys using Time Series Methods. *Journal of the American Statistical Association*, 69, pp. 674-678.

Tiller, R.B. (1992). Time Series Modelling of Sample Survey Data from the U.S. Current Population Survey. *Journal of Official Statistics*, vol. 8, pp. 149-166.

Appendix

A.1 Diagnostics for the prediction errors

BIC and AIC are popular criteria for comparing and selecting statistical models. The use of these criteria in the context of structural time series models is not straightforward and is therefore not discussed and applied in this paper. See Boonstra et al. (2008) for a more detailed discussion about the use of AIC and BIC in this context.

The one-step forecast errors or prediction errors $\upsilon_{t,d}$ are defined as the difference between the one-step forecasts $\theta_{t|t-1,d}$ and the GREG estimates. The underlying assumptions of the state-space model are that the disturbances of the measurement equation and the transition equation are normally distributed and serially independent with constant variance. Under this assumption, the standardized prediction errors $e_{t,d}$ are also normally and independently distributed with constant variance. To check the properties of the standardized prediction errors $e_{t,d}$, standard tests known from time series literature are applied (see e.g. Durbin and Koopman, 2001). A short description of how the tests are applied in this paper is given below:

The model estimates for the first two years are excluded from the model evaluation because the model estimates from the first periods cannot be used due to the diffuse initialisation of the model. All tests are done at a 5% significance level.

The standardized prediction errors are computed by

$$e_{t,d} = \frac{\upsilon_{t,d}}{\sqrt{\mathrm{var}(\theta_{t|t-1,d}) + \sigma^2_{v,d} Var_{t,d}}} . \qquad \text{(A.1)}$$

For multivariate models, this is a simplified way for the standardization, because a possible covariance between the one-step forecasts $\theta_{t\,t-\;,d}$ of the different domains is ignored.

It is tested for normality by computing skewness and kurtosis and testing whether these values are significantly different from the values of the normal distribution.

Under the assumption of normally and independently distributed disturbances, the Kalman filter yields best unbiased estimators for the state variables. If the normality assumption is not met, the Kalman filter still yields the best linear unbiased estimates for the state variables. Therefore, a small violation of the normality assumption is no reason to reject a model. The test is still useful, for example because the violation can be caused by an outlier. Including this outlier in the model can be an improvement.

With the F-test for heteroscedasticity can be checked whether the standardized prediction errors are distributed with equal variance. Generally, approximately the first third of the time series is compared with the last third. When the data until December 2007 are used,

$$H_d = \frac{\sum_{t=97}^{144} e_{t,d}^2}{\sum_{t=25}^{72} e_{t,d}^2}$$

is computed.

When the data until December 2008 are used,

$$H_d = \frac{\sum_{t=109}^{156} e_{t,d}^2}{\sum_{t=25}^{72} e_{t,d}^2}$$

is computed.

Under the hypothesis of homoscedasticity, H_d is $F_{48,48}$ -distributed (Durbin and Koopman, 2001).

By computing the sample autocorrelation function, serial independence of the standardized prediction errors can be tested for. The sample autocorrelation function is defined as

$$\hat{\rho}(h) = \frac{\hat{\gamma}(h)}{\hat{\gamma}(0)},$$

where $\hat{\gamma}(h) = \frac{1}{T-24} \sum_{t=25}^{T-h} (e_{t+h} - \bar{e})(e_t - \bar{e})$

and $\bar{e} = \frac{\sum_{t=25}^{T} e_t}{T-24}$.

The sample autocorrelation function can be computed for $h = 1,2,...,T-25$, but for large h it is based on very little information and the estimates will become unstable. Following the advice of Brockwell and Davis (2002), to compute the sample autocorrelation function for $h < {T-24}/{4}$, here the sample autocorrelations are computed for $h < 27$. The sample autocorrelations of white noise are approximately normally distributed with expectation zero and variance $1/n$, where *n* is the length of the series. This means that no more than around 5% of the sample autocorrelations of the prediction errors should be outside the bounds $1.96/\sqrt{120} = 0.179$.

The sample autocorrelations can be plotted in a graph to see whether there are some patterns visible, for example cyclical patterns.

A.2 Cross-validation using the prediction error

In time series analysis, cross-validation can be used to asses the predictive power of a model. The mean of the prediction errors, of the absolute values of the prediction errors or the mean of the squared prediction errors can be used as a form of cross-validation to measure the deviation of the model forecasts from the GREG estimates. For the cross-validation, one step forcasts $\vartheta_{t|t-1,d}$ are computed with hyperparameter estimates based on the information available up to *t*-1. Then the prediction errors $\tau_{t,d}$ are defined as the difference between the one-step forecasts $\vartheta_{t\,t-\,,d}$ and the GREG estimates.

In the evaluation of the models with data until December 2007, the predictions of last two years are used. The mean of the prediction errors is computed as

$$MPE_d = \frac{1}{24}\sum_{t=121}^{144}\tau_{t,d} \ . \qquad \text{(A.2)}$$

A large absolute value indicates that the model estimates are biased.

The mean absolute prediction error is computed as

$$MAPE_d = \frac{1}{24}\sum_{t=121}^{144}\left|\tau_{t,d}\right| \qquad \text{(A.3)}$$

and the root of the mean squared prediction error is

$$RMSPE_d = \sqrt{\frac{1}{24}\sum_{t=121}^{144}\tau_{t,d}^2} \ . \qquad \text{(A.4)}$$

A small value for MAPE and RMSPE implies that the model predicts the true population value well.

In the second evaluation, where the data of 2008 is added, only the data of the last year is used, thus

$$MPE_d = \frac{1}{12}\sum_{t=145}^{156}\tau_{t,d}$$

and similar for MAPE and MSPE.

It is also possible to compute weighted *MPE*, *MAPE*, and *RMSPE*. Possible weights are the inverse of the variances or standard errors of the GREG estimates. Because these variances are quite similar in the considered period and because they are estimated with some uncertainty, this would be at most a small improvement in this application.

Appendix A.3: Figures

Figure A.1: GREG estimates and filtered estimates under model 1 for domain 2

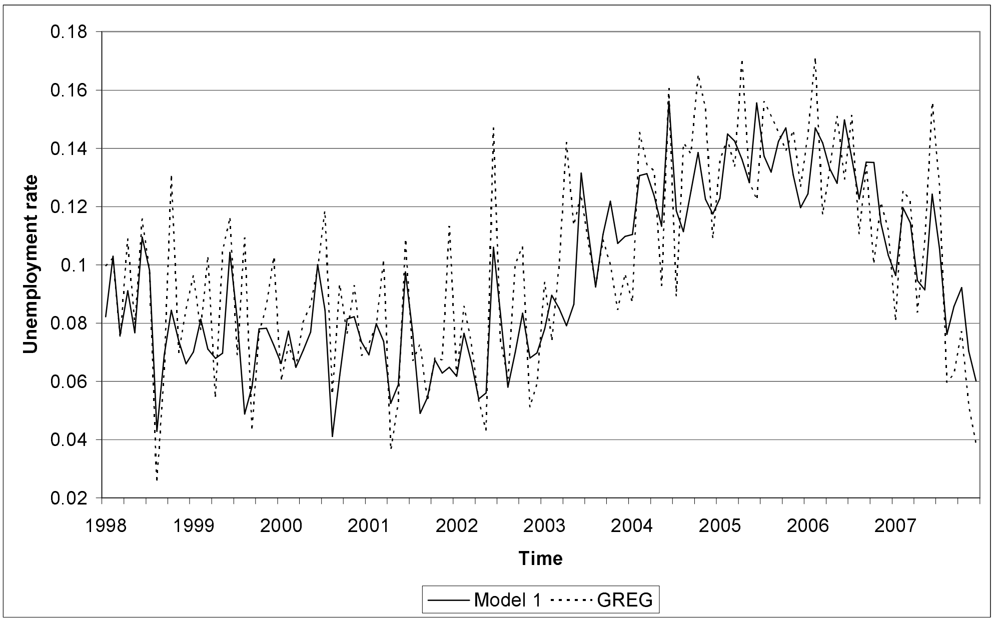


Figure A.2: GREG estimates and filtered estimates under model 1 for domain 3

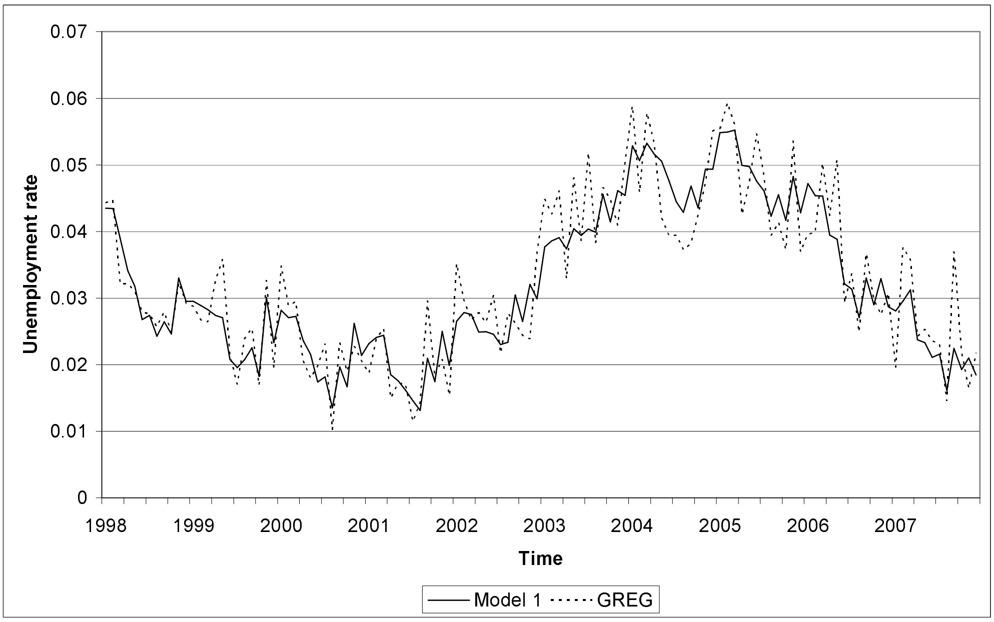


Figure A.3: GREG estimates and filtered estimates under model 1 for domain 4

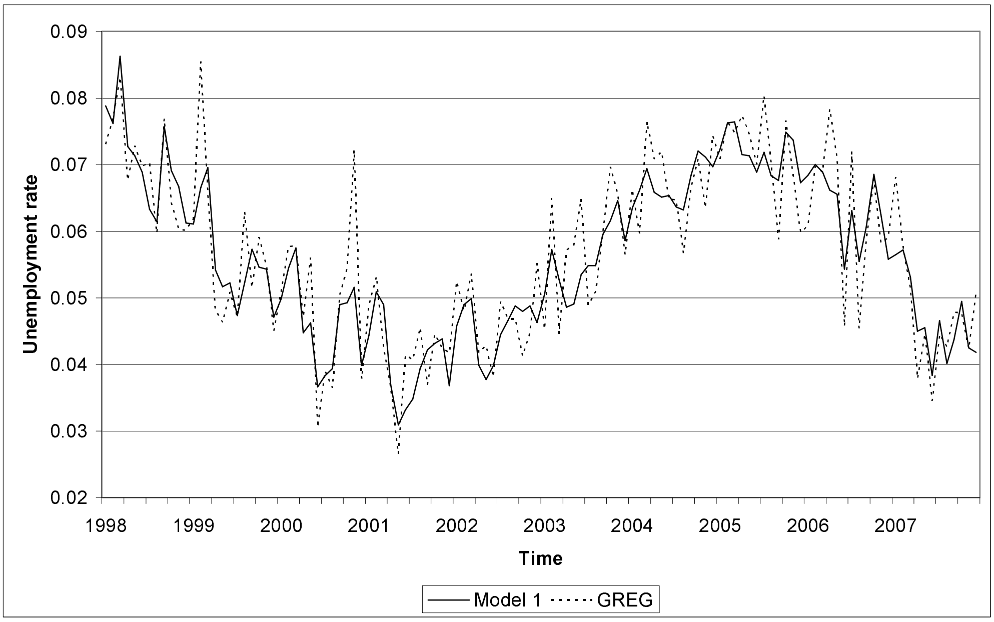


Figure A.4: GREG estimates and filtered estimates under model 1 for domain 5

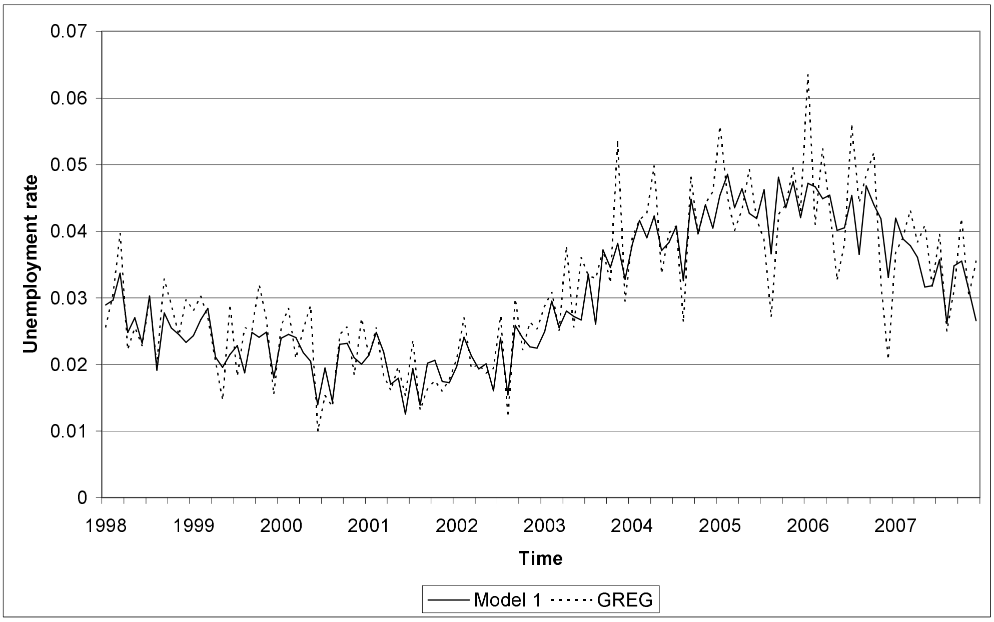


Figure A.5: GREG estimates and filtered estimates under model 1 for domain 6

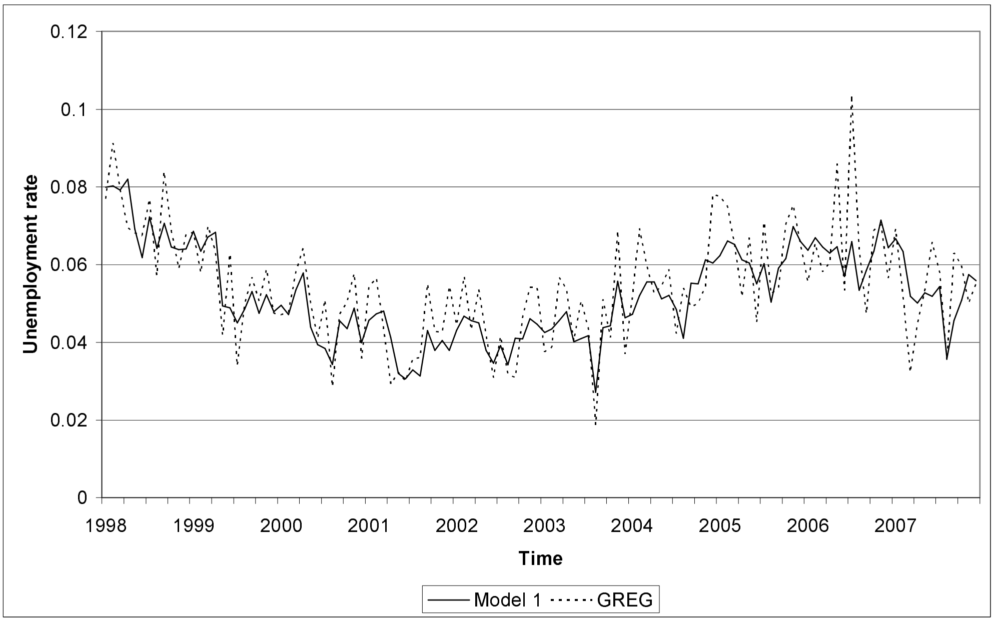


Figure A.6: Filtered estimates under model 1 and model 2 for domain 2

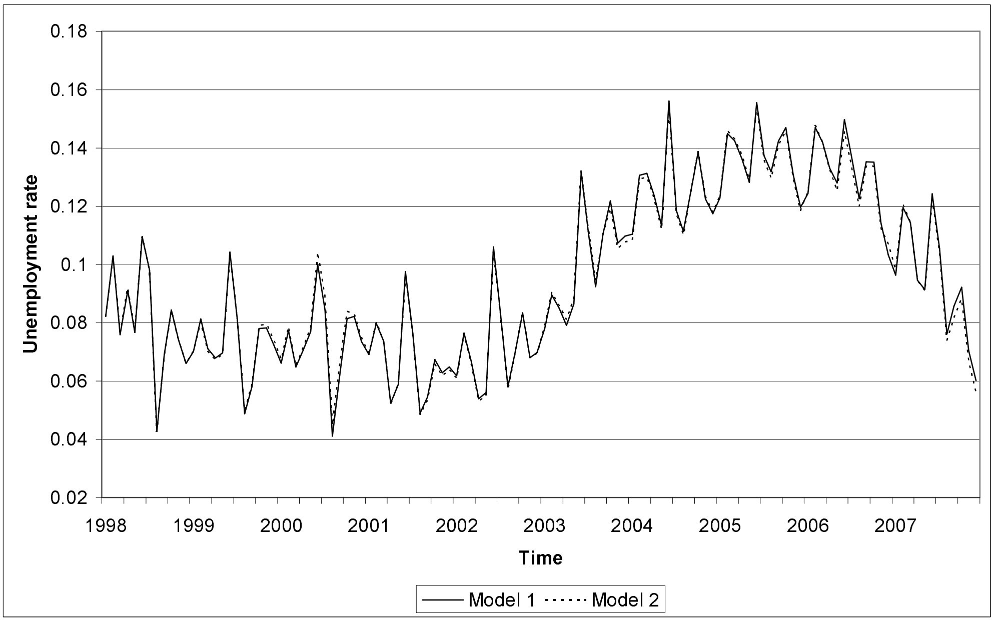


Figure A.7: Filtered estimates under model 1 and model 2 for domain 3

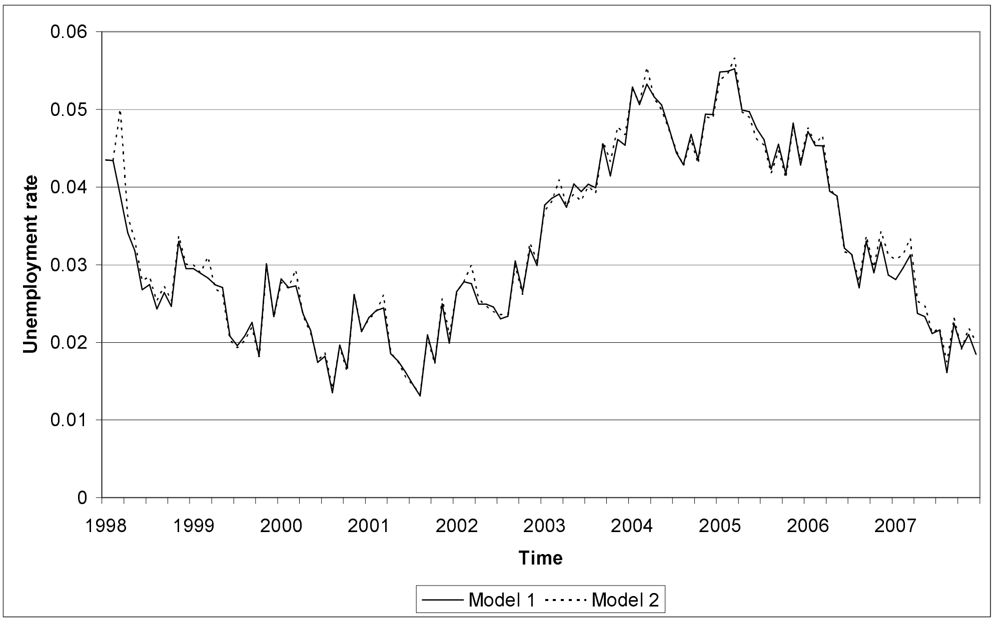


Figure A.8: Filtered estimates under model 1 and model 2 for domain 4

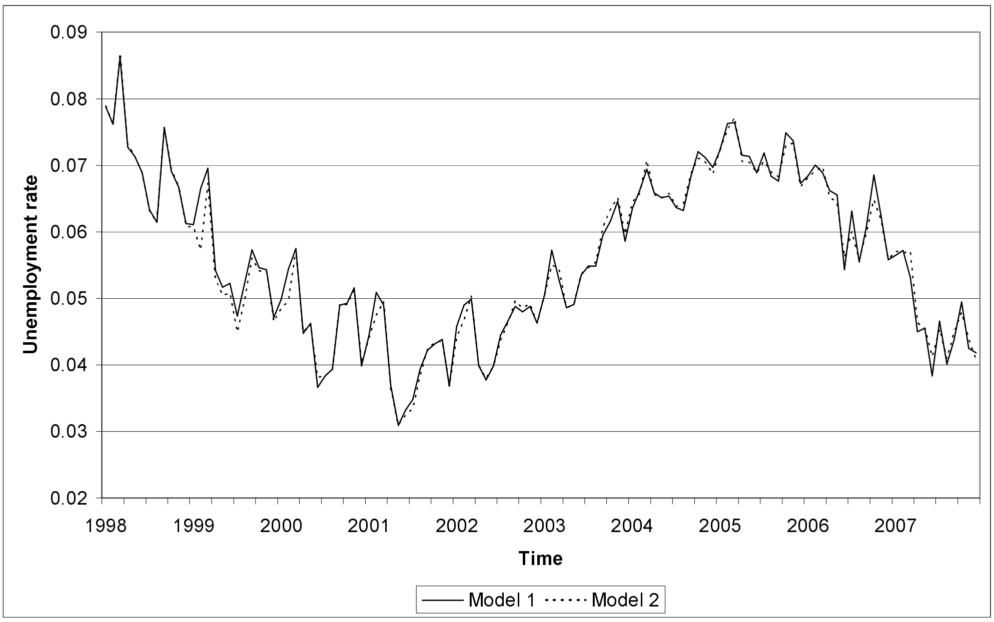


Figure A.9: Filtered estimates under model 1 and model 2 for domain 5

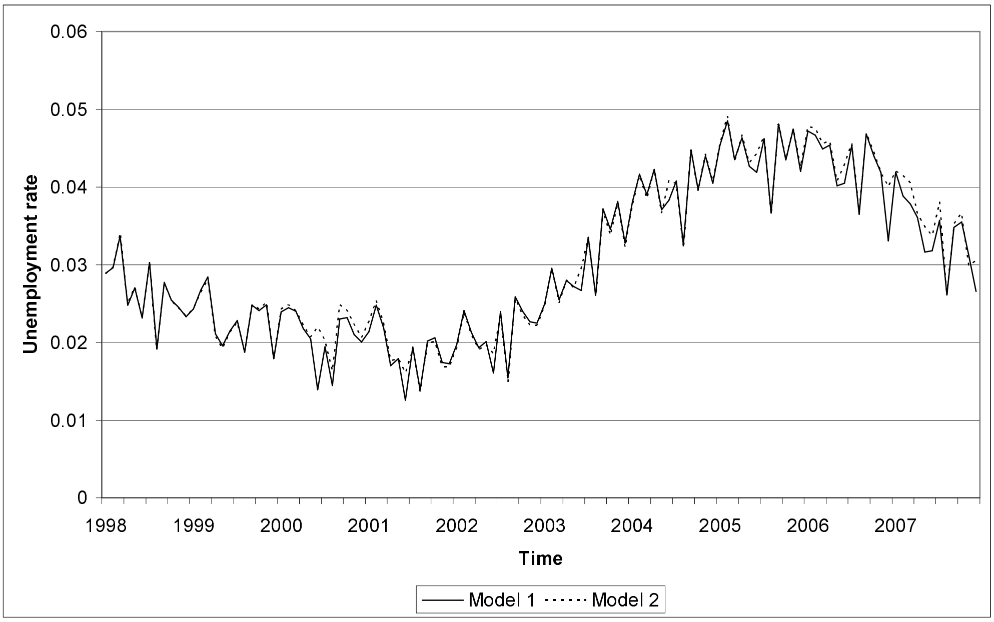


Figure A.10: Filtered estimates under model 1 and model 2 for domain 6

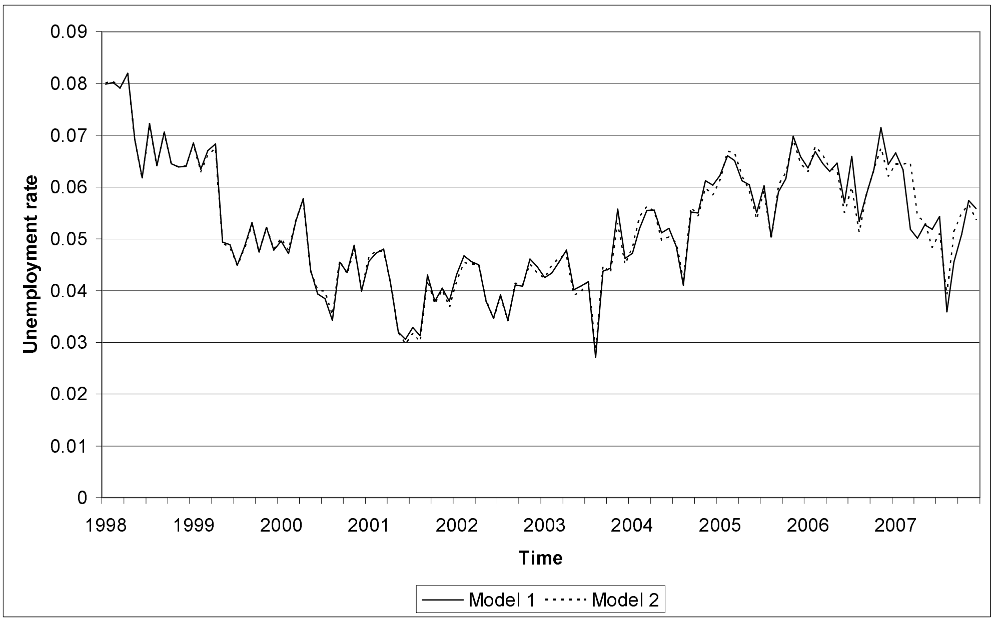


Figure A.11: Filtered estimates under model 2 and model 3 for domain 2

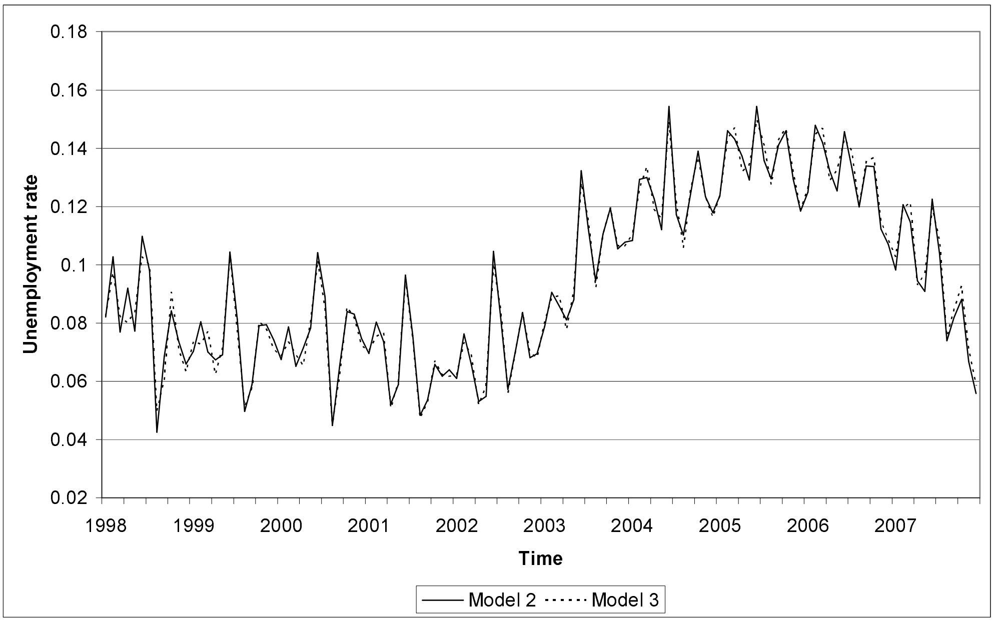


Figure A.12: Filtered estimates under model 2 and model 3 for domain 3

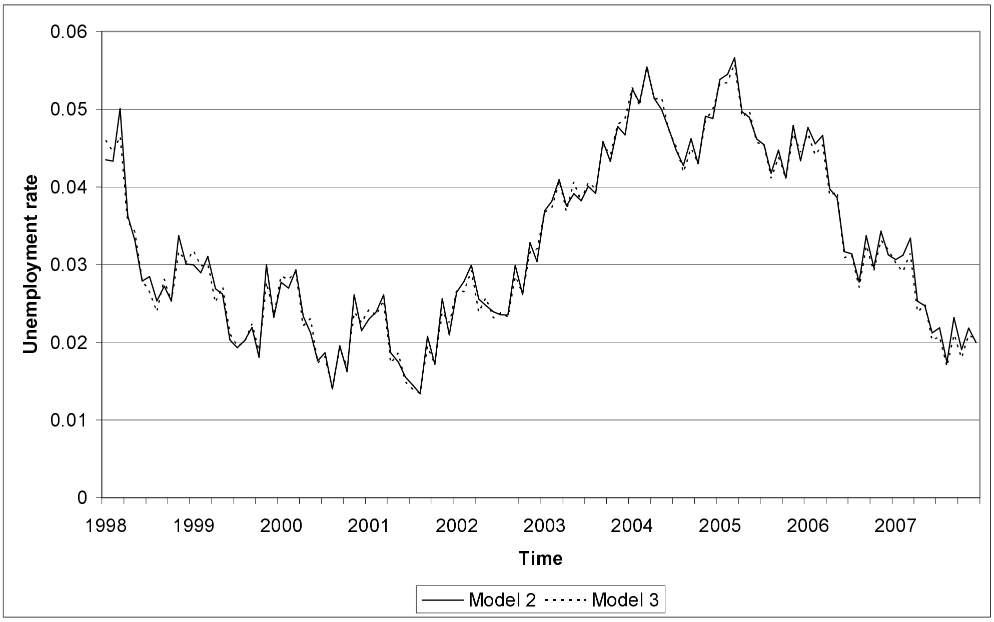


Figure A.13: Filtered estimates under model 2 and model 3 for domain 5

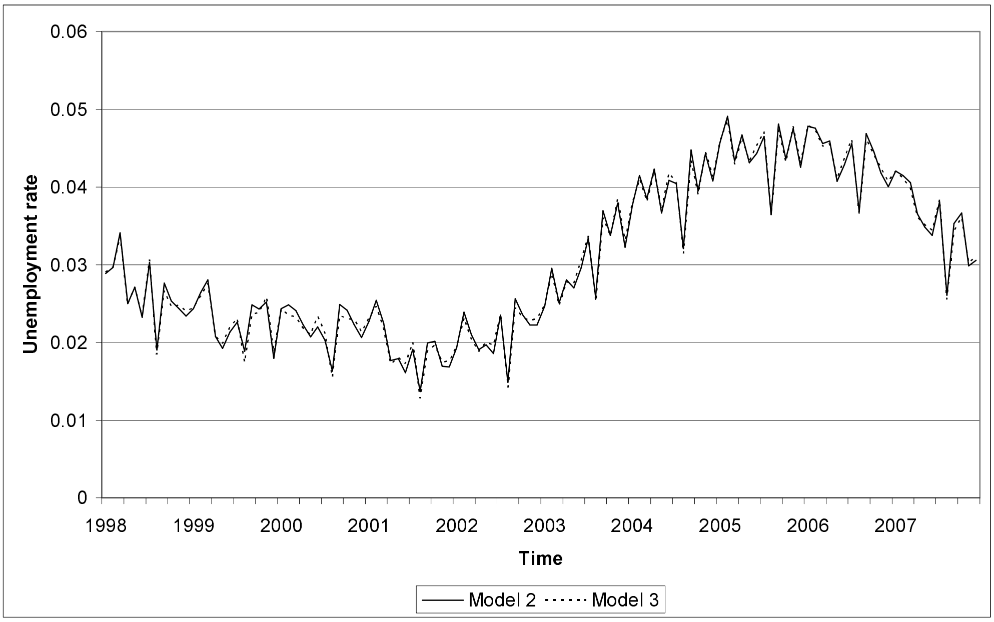


Figure A.14: Filtered estimates under model 2 and model 3 for domain 6

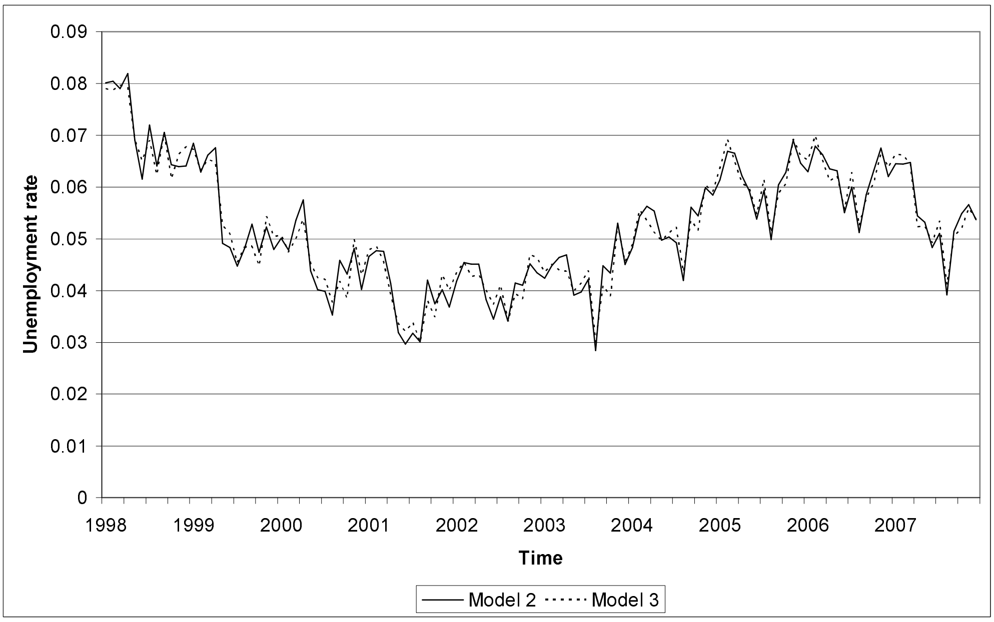


Figure A.15: Filtered estimates under model 1, model 4, and model 6 for domain 2

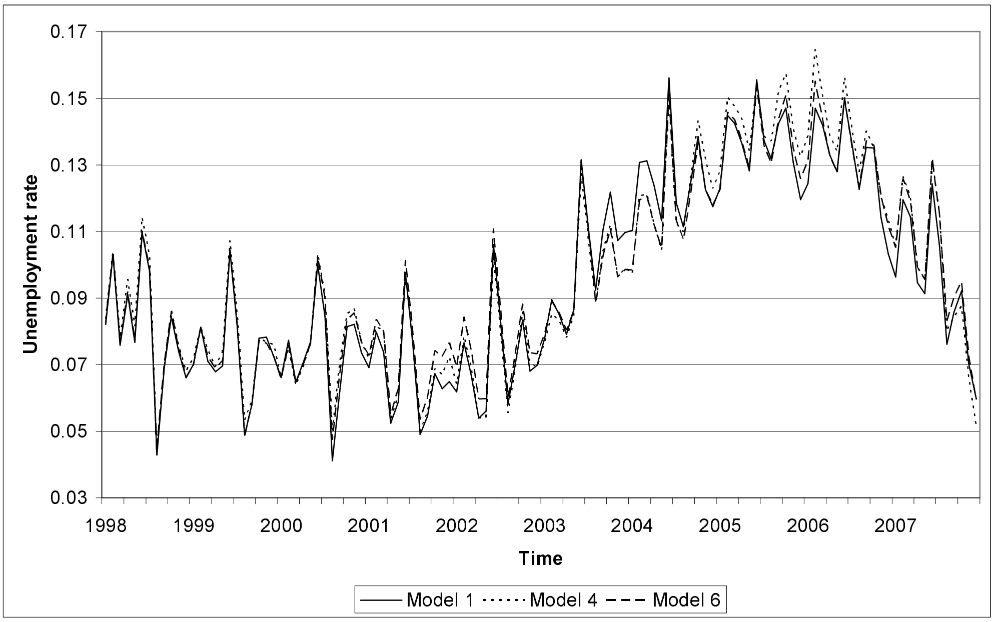


Figure A.16: Filtered estimates under model 1, model 4, and model 6 for domain 3

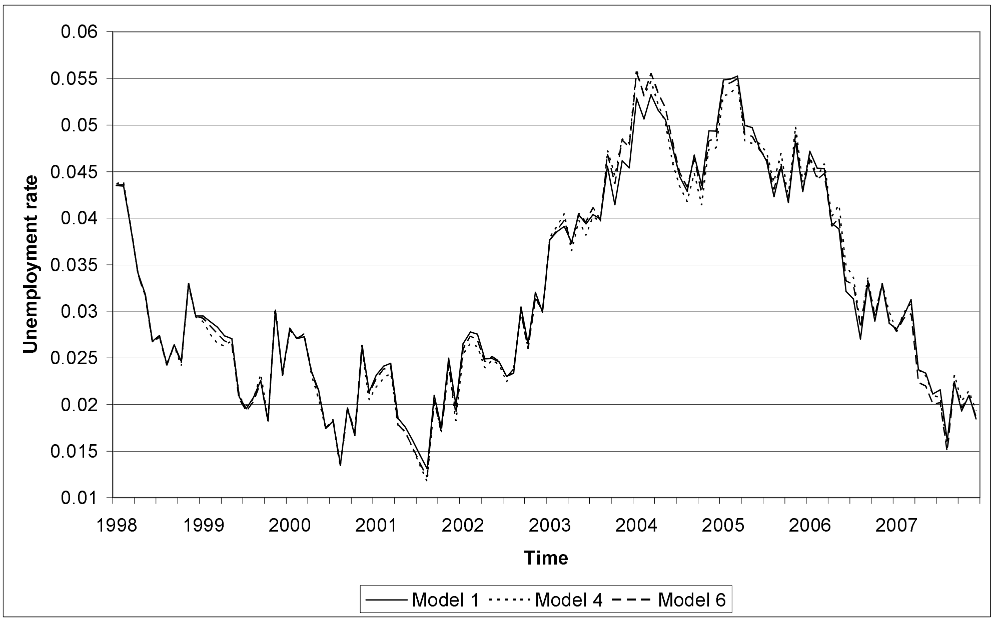


Figure A.17: Filtered estimates under model 1, model 4, and model 6 for domain 4

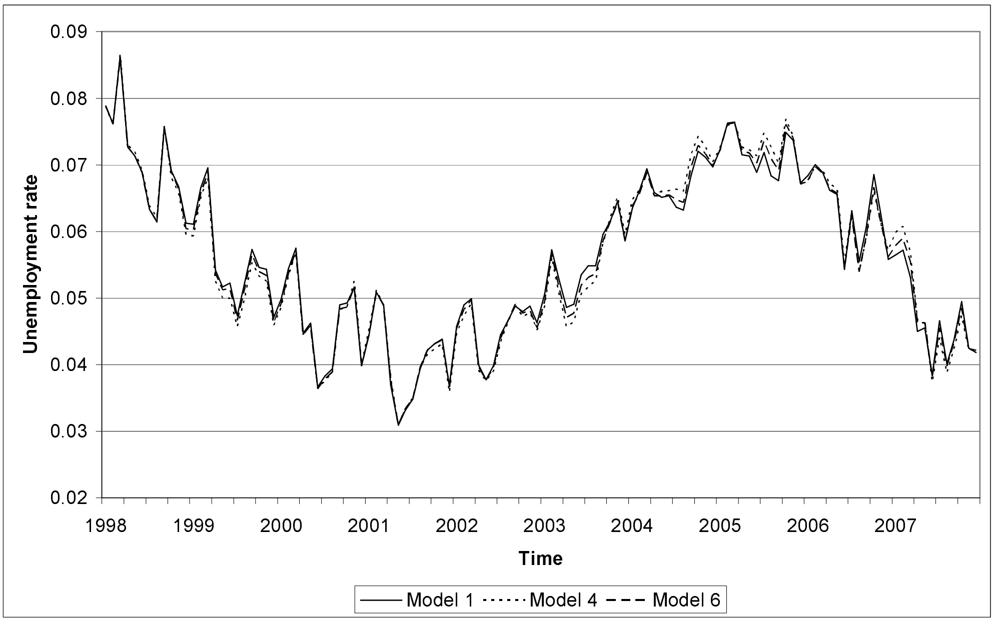


Figure A.18: Filtered estimates under model 1, model 4, and model 6 for domain 5

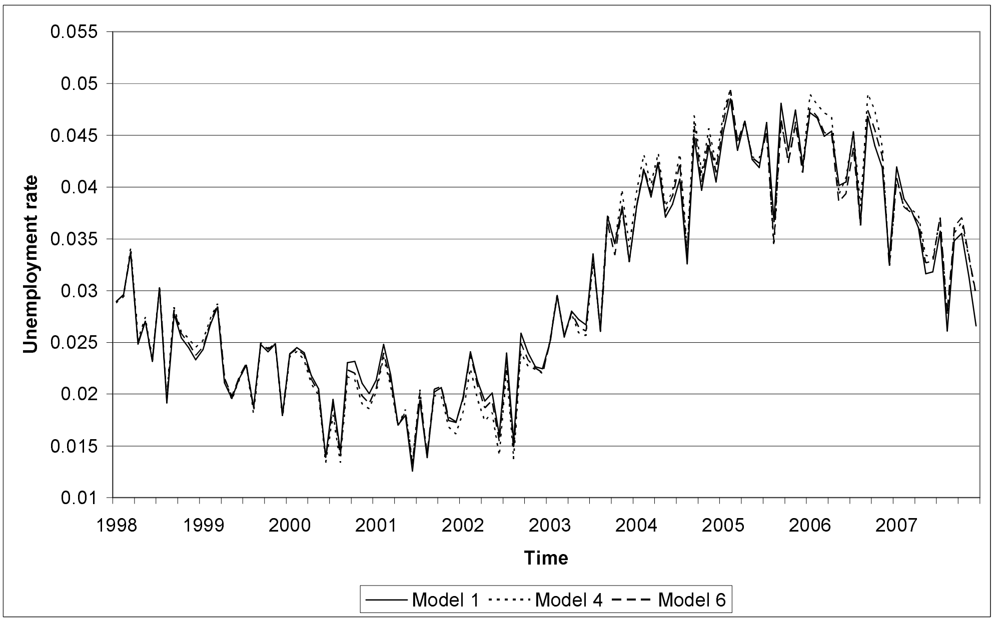


Figure A.19: Filtered estimates under model 1, model 4, and model 6 for domain 6

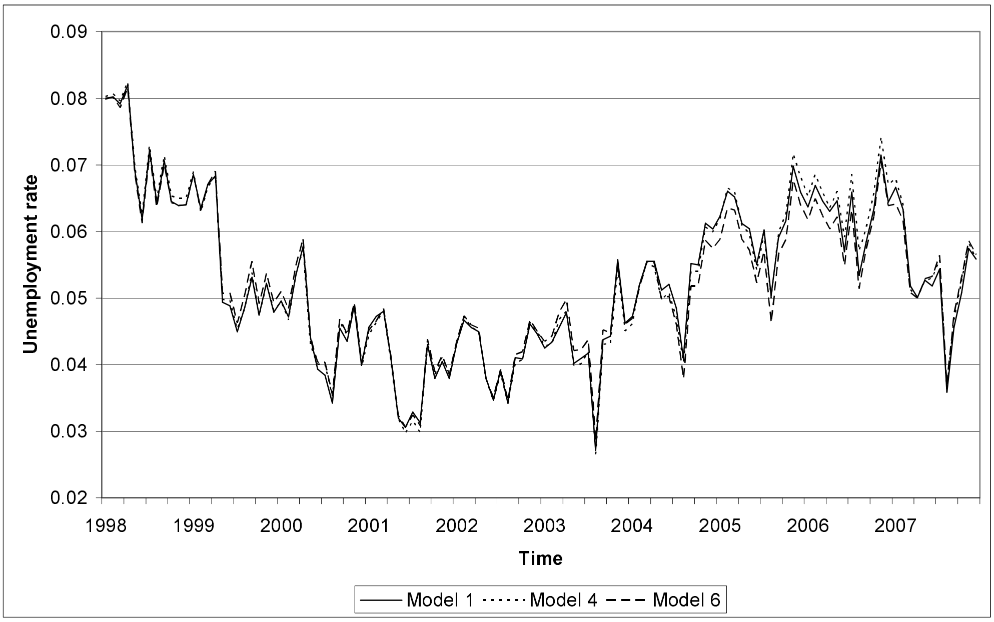


Figure A.20: Filtered estimates under model 3, model 5, and model 7 for domain 2

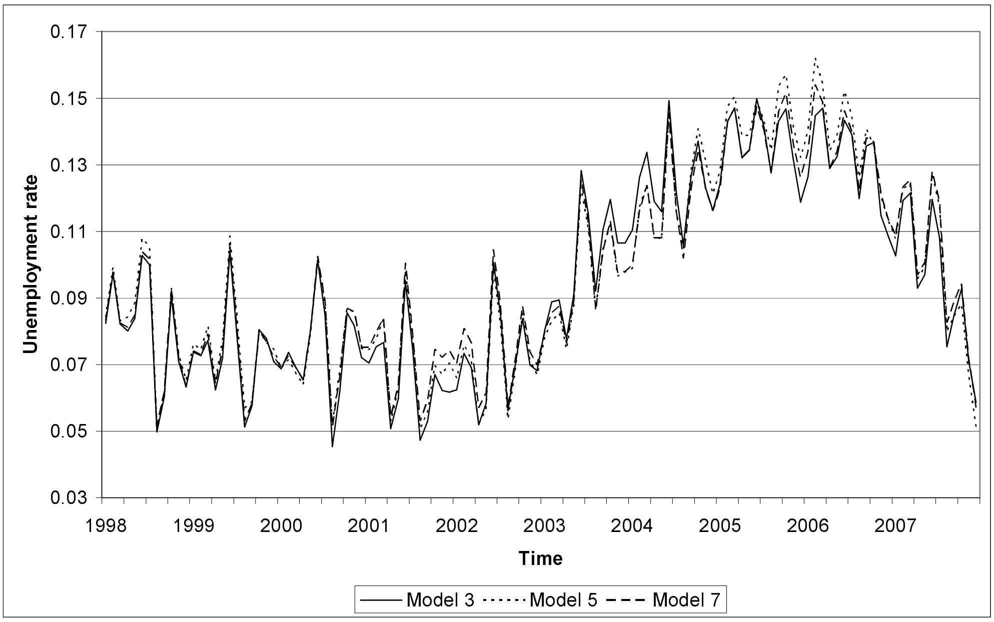


Figure A.21: Filtered estimates under model 3, model 5, and model 7 for domain 3

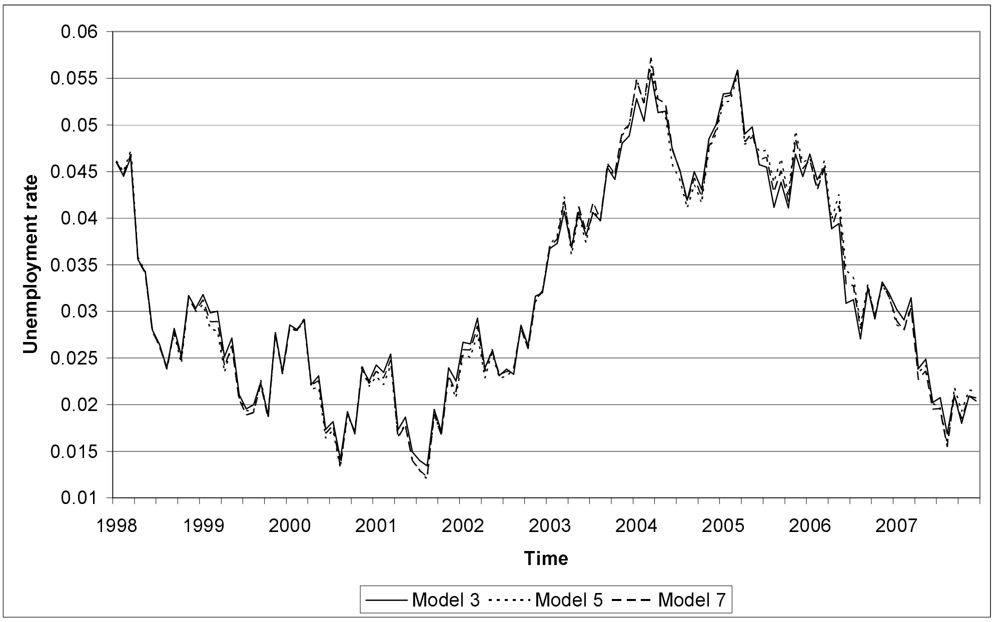


Figure A.22: Filtered estimates under model 3, model 5, and model 7 for domain 4

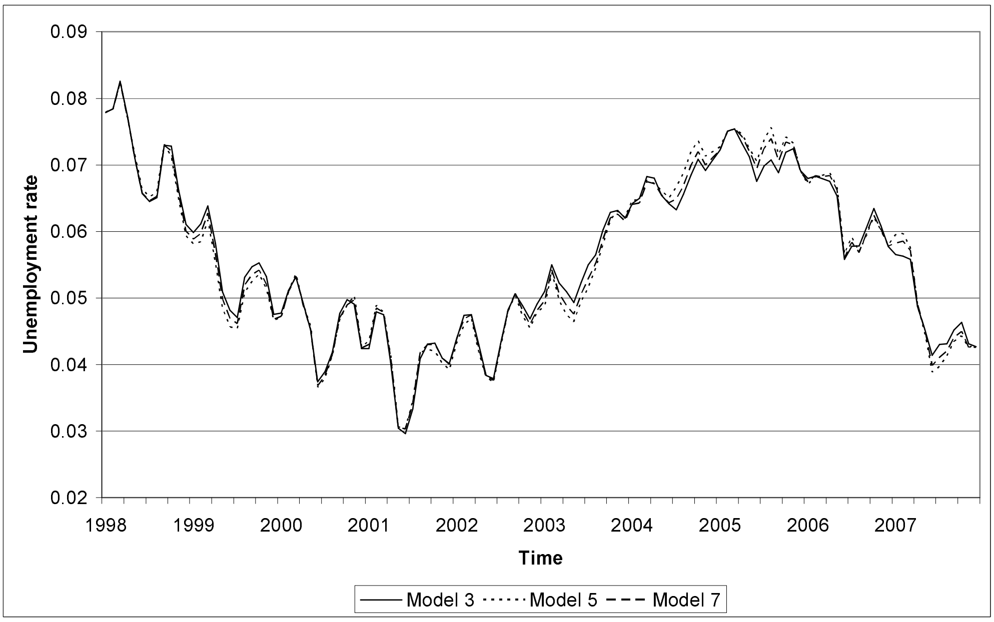


Figure A.23: Filtered estimates under model 3, model 5, and model 7 for domain 5

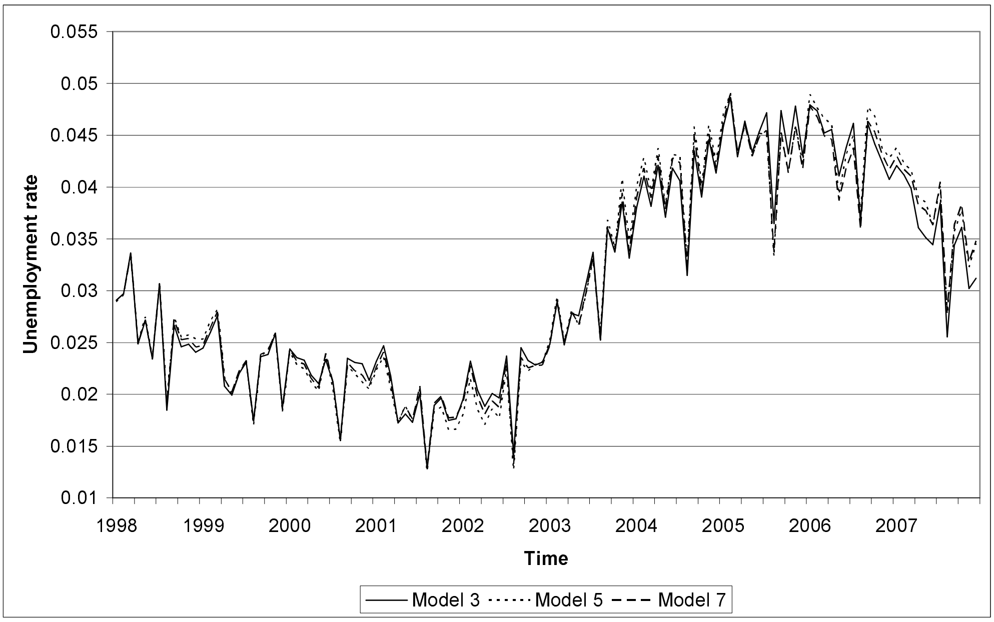


Figure A.24: Filtered estimates under model 3, model 5, and model 7 for domain 6

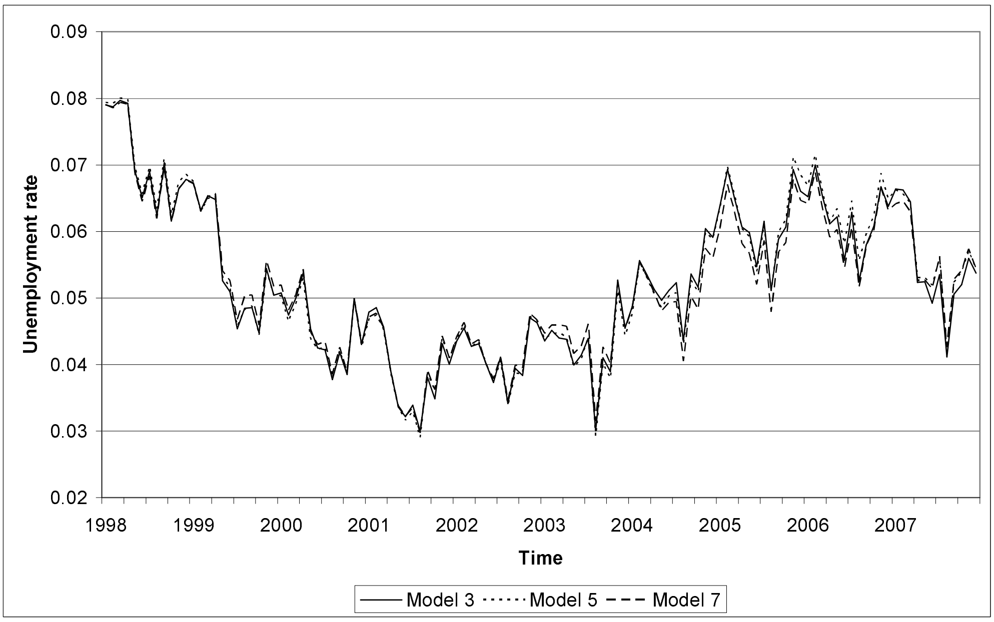


Figure A.25: Smoothed estimates of seasonal effect of domain 2 under model 3

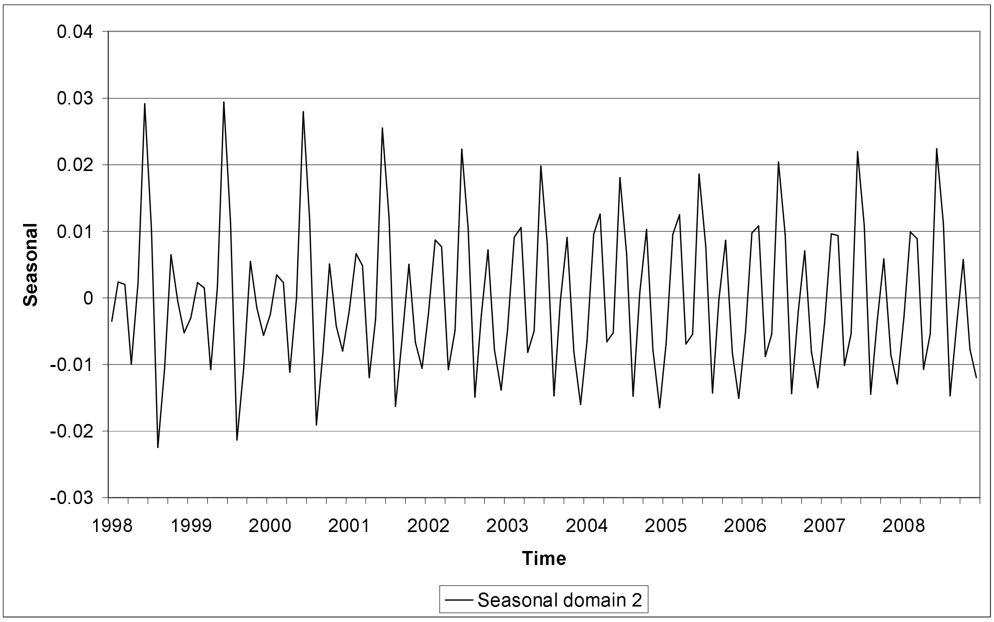


Figure A.26: Smoothed estimates of seasonal effect of domain 3 under model 3

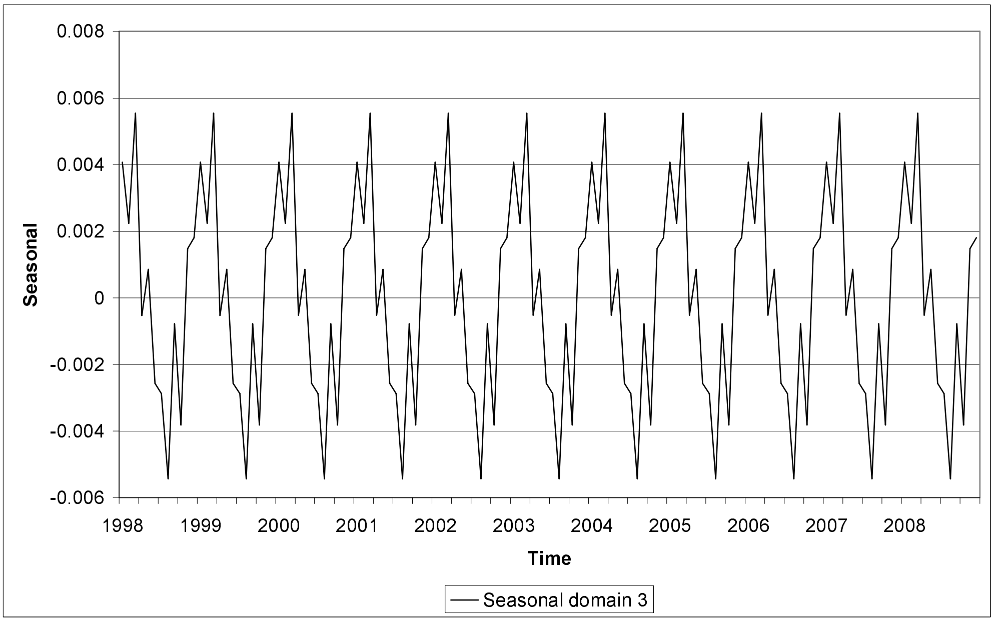


Figure A.27: Smoothed estimates of seasonal effect of domain 4 under model 3

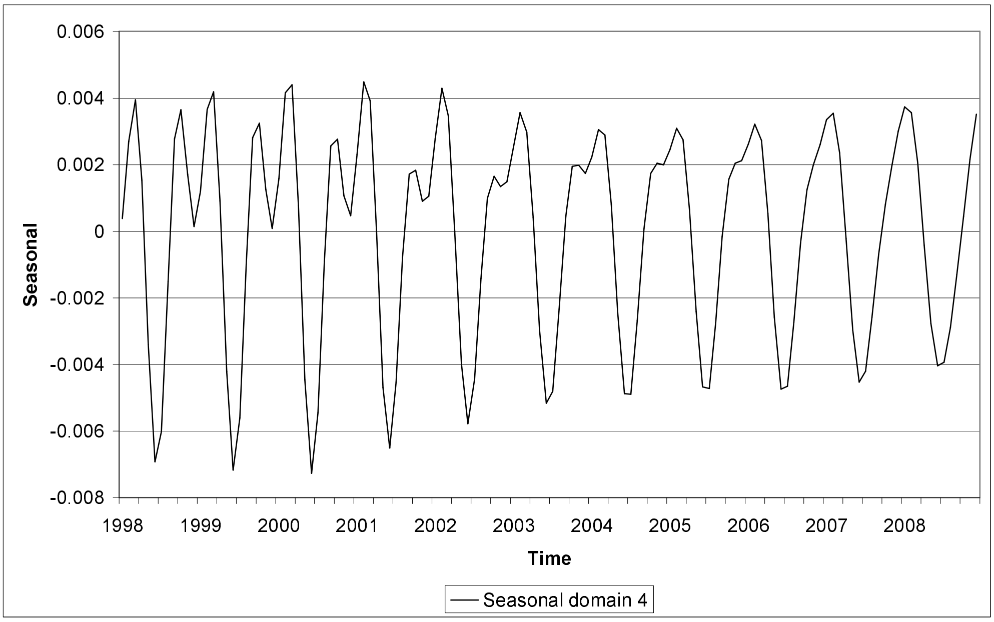


Figure A.28: Smoothed estimates of seasonal effect of domain 5 under model 3

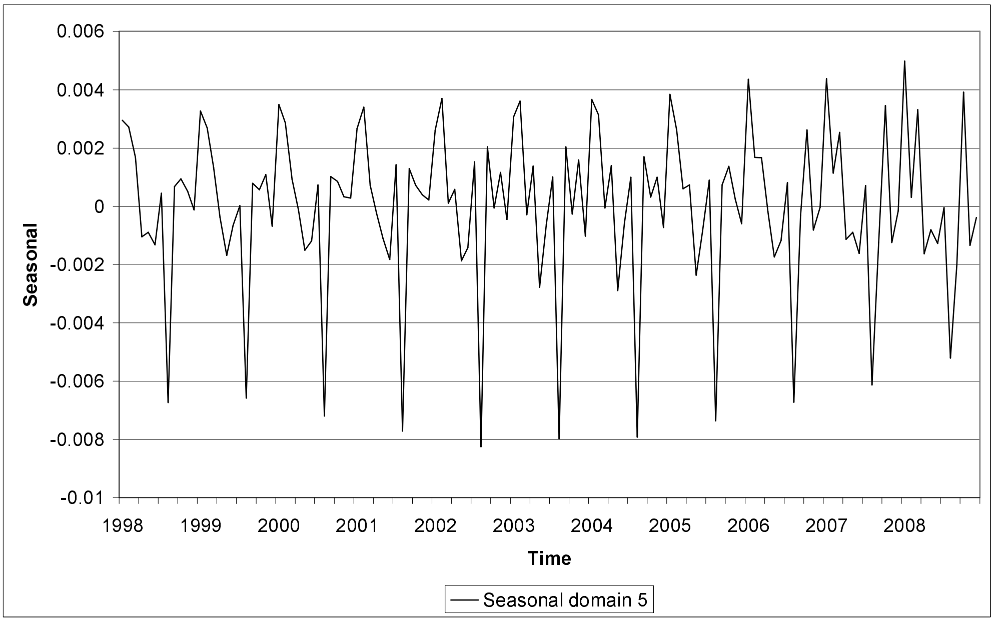


Figure A.29: Smoothed estimates of seasonal effect of domain 6 under model 3