ESTIMATING MONTHLY GDP IN A GENERAL KALMAN FILTER FRAMEWORK: EVIDENCE FROM SWITZERLAND.[*,†]

NICOLAS A. CUCHE[‡] MARTIN K. HESS[§]

First draft: Oct. 15, 1998
This version: Mar. 25, 1999

Abstract

In this paper, we estimate deseasonalized monthly series for Swiss gross domestic product at constant prices of 1990 for the period 1980-1998. They are consistent with the quarterly figures estimated by the Federal Office for Economic Development and Labour and are obtained by including information contained in related series. We present a general approach using the Kalman filter technique nesting a great variety of interpolation setups. We evaluate competing models and provide a time series that can be used by other researchers.

JEL codes: E32, E37
Keywords: Interpolation, Kalman filter, National accounting.

[*]We thank Ramses Abul Naga, Philippe Bacchetta, Harris Dellas, Giovanni Leonardo, Iwan Meier, Jeffrey Nilsen, Walter Wasserfallen, Mark Watson, and seminar participants at the University of Lausanne for their helpful comments and suggestions. We have also benefitted from discussions with experts from the Federal Office for Economic Development and Labour, especially Bruno Parnisari, and the Federal Statistics Office. All remaining errors are ours.

[†]For correspondence: Studienzentrum Gerzensee, Postfach 21, CH-3115 Gerzensee, Switzerland. Email: ncuche@szgerzensee.ch, mhess@szgerzensee.ch.

[‡]University of Lausanne and Studienzentrum Gerzensee.

[§]University of Bern and Studienzentrum Gerzensee.

1 Introduction

For economic studies using quarterly data, a low number of observations can cause serious flaws in the quality of quantitative analysis. In vector autoregressions (VAR) with relatively short time series for example, many degrees of freedom are used up in the estimates, reducing drastically the power of the estimation. Moreover, monthly frequency is sometimes implied by the assumptions of the model being estimated, while only quarterly data are released. Estimates of Swiss gross domestic product (GDP) for every quarter are dating back to 1965[1].

Therefore, economists are sometimes forced to use variables that proxy GDP and that are available at a higher frequency. A common proxy in many countries is industrial production (IP) which is often recorded monthly and which comoves closely with GDP. In Switzerland, it is difficult to find such a monthly indicator for aggregate productive activity. The IP index is a series at a quarterly frequency, and other series like business surveys or filled orders can only be used as GDP proxies with strong reservations on the interpretation. Hence, in cases where adequate proxies are not at hand, monthly estimates of GDP by interpolation[2] can provide a solution to this problem.

Whether to replace a proxy variable by an interpolated one or not depends on the available data series and on the empirical economic model considered. The evaluation of the trade-off between the potential benefits and disadvantages of both approaches is beyond the scope of this paper and is omitted. The goal of this paper is to provide a monthly deseasonalized[3] real GDP series for empirical research.

Chow and Lin [1971] were the first to present a coherent and easily applicable econometric approach that handles interpolation problems for stock and flow variables. Assuming a linear relation between the series of interest (series for which observations are missing, i.e. monthly GDP) and other data with more frequent recording (related series), they estimate a univariate regression equation. This multiple regression approach is flexible enough to take into account heteroskedasticity and autocorrelation in the residuals.

More recent approaches make use of the Kalman filter (Harvey and Pierse

[1]The official figures concerning the quarterly GDP are published by the Federal Office for Economic Development and Labour. Furthermore, an official annual GDP is calculated by the Federal Statistics Office producing the national income accounts. The quarterly estimates are then corrected and published again to match the official annual statistics.

[2]We understand interpolation as a process of computing flow or stock series at a higher frequency than the original one. In this terminology we do not distinguish between interpolation and distribution which is often done in studies with both, stock and flow variables. Here, we present models that serve exclusively for interpolation and not for out-of-the-sample predictions.

[3]In our view, deseasonalized time series are of greater interest as they are handy to use in economic models. To estimate a seasonalized series, the seasonality would have to be estimated separately and then added to the deseasonalized series, as done for example by the Federal Office of Economic Development and Labour for quarterly GDP estimates.

[1984], and Bernanke, Gertler and Watson [1997]). This dynamic framework is much more flexible, since it is capable of nesting various models and is more promising because the first estimates are updated as new information arrives.

In this paper, the focus is directed on econometric details such as the issue of stationarity and cointegration in different Kalman filter configurations. Recent innovative techniques are analyzed theoretically and then evaluated empirically. An overview of estimated monthly GDP series produced by various model setups is provided. We evaluate different combinations of methods and related series with the aim to get the most appropriate monthly GDP. For this task, several selection criteria as well as a simulated interpolation from annual to quarterly data are used.

Before estimating the model, we evaluate competing related series. We identify the series containing the highest amount of information for the interpolation. The choice criteria for the related monthly series are based on the expenditure definition of GDP and on statistical properties of the comovement with GDP. However, the dearth of Swiss data at higher frequency limits severely the choice of these variables. Therefore, we consider other related series, for example, foreign aggregate economic activity, as alternatives for interpolation. In fact, all related series that closely and robustly move together with quarterly GDP could be appropriate series helping to extract monthly GDP. With these related series available, it is then possible to estimate monthly GDP for Switzerland for 1980-1998[4] in different model setups.

The paper is organized as follows. It starts in Section 2 with a short survey of the interpolation literature. In Section 3, we briefly review the Kalman filter methodology. The different interpolation models are presented. In Section 4, various related series are evaluated and described. An overlook of the results is given in Section 5. We then evaluate the appropriateness of these interpolations. Section 6 concludes.

2 Related Literature

As Lanning [1986] illustrates, economists facing missing data have basically two different ways to solve that problem. A first approach is to estimate the missing data simultaneously with the model parameters, thereby considering the missing observations as any other parameters. The second way is a two-step approach where in a first step the missing data, which could be independent from the

[4]We exclusively concentrate our investigation on the period 1980-1998 because these figures are compatible with the new national accounting system in Switzerland, the European System of National Accounting (ESNA) 78. In Switzerland it was introduced in 1996, but the Federal Office for Economic Development and Labour calculated quarterly GDP figures back to 1980. See Schwaller and Parnisari [1997] for a very good survey. The structural break is too important between the former and new system for not considering it.

economist's model, are interpolated. In a second step, the new augmented series are used to estimate the economist's model. Lanning found that the simultaneous approach yields estimates of the economist's model parameters that have a greater variance and thus are less reliable than the model parameters estimated with complete data in the second stage. Based on these empirical findings, he suggests the use of the two-step approach. Related literature on the latter procedure can be subdivided in the following three classes[5].

First, the seminal approach for the use of the univariate multiple regression technique with related series was established by Chow and Lin [1971] and [1976] who presented a unified framework which allows to treat the interpolation of stocks and flows variables. This approach was able to overcome the problems faced by Friedman [1962] who treated stocks and flows in different ways. Specifically, they could deal with the requirement that if an observed flow value is distributed among the corresponding subintervals, the higher frequency estimates must add up to the observation of the original lower frequency variable. Until now, this static regression approach has been widely used for interpolation due to its easy implementation compared to the state-space approach. This argument seems to more than just outweigh the potential advantages of more sophisticated procedures like the Kalman filter. An annual GDP was for example interpolated for Mexico by De Alba [1990]. Schmidt [1986] gives a good survey of this method interpolating personal income of US regions.

Second, still with related series, Denton [1971], Fernandez [1981], and Litterman [1983] proposed an approach that minimizes a weighted quadratic loss function on the difference between the series to be estimated and a linear combination of the observed related series. This strategy nests the Chow and Lin regression, but allows for more complicated assumptions about the driving process of the interpolated variable and the use of data in first difference. An illustration with Portuguese data is given in Pinheiro and Coimbra [1993].

Third, Bernanke, Gertler and Watson [1997] have recently used a state-space model to interpolate real GDP in the US. Their approach is to first estimate monthly components of nominal GDP and the GDP deflator and then to aggregate the individual estimates. The lack of data in Switzerland prevents us from disaggregating GDP for our interpolation. The methodology they followed was suggested by Harvey and Pierse [1984] who provide a general framework - state-space formulations for stock and flow variables, and for stationary and non-stationary series, with and without related series - to estimate missing observations in economic time series. Solving such state-space models requires the use of the Kalman filter. A Kalman filter interpolation has also been done for Canadian GDP by Guay et al. [1990].

[5]The signal extraction literature is very vast and difficult to objectively classify. Here, we only review the interpolation literature, without considering general approaches such as the problem of unobserved components in economic time series or the estimation of irregularly missing data.

Hereafter, we present a state-space framework introduced by Harvey and Pierse [1984]. This formulation allows us to rewrite models using related series following Chow and Lin [1971], Denton [1971], Fernandez [1981], Litterman [1983], Bernanke, Gertler and Watson [1997], and simpler models that do not use related series.

3 Models

3.1 Kalman Filter

To extract unobserved variables is a common problem in economics. One useful method for extracting signals is to write down a model linking the unobserved and observed variables in a state-space representation according to Kalman [1960] and [1963]. The multivariate Kalman filter is an algorithm for sequentially updating a linear projection on the vector of interest. A general review is given here and a more detailed description in the appendix[6]. We present the various configurations of the state-space system we use in the next section on interpolation models.

The state-space representation is given by a system of two vector equations. First, the state or transition equation describes the dynamics of the state vector ($\boldsymbol{\xi}_t$) containing the unobserved variables we want to estimate. The second type of equation represents the observation or measurement equation linking the state vector to the vector containing the observed variables $\left(\mathbf{y}_t^+\right)$. The equations of this system for $t = 1, \ldots, T$ where T is the number of monthly observations are the following:

$$\begin{aligned} \boldsymbol{\xi}_{t+1} &= \mathbf{F}_t\boldsymbol{\xi}_t + \mathbf{C}_t'\mathbf{x}_{t+1} + \mathbf{R}_t\mathbf{v}_{t+1} \qquad (1) \\ \mathbf{y}_t^+ &= \mathbf{A}_t'\mathbf{x}_t^* + \mathbf{H}_t'\cdot\boldsymbol{\xi}_t + \mathbf{N}_t\mathbf{w}_t \qquad (2) \end{aligned}$$

In addition to the unobserved and the observed variables of interest, vector equations (1) and (2) contain the so-called related series ($\mathbf{x}_t$) and ($\mathbf{x}_t^*$) as exogenous variables in each equation. Both equations have error terms multinormally distributed: $\begin{pmatrix} \mathbf{v}_t \\ \mathbf{w}_t \end{pmatrix} \sim N\left(\begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix}, \begin{pmatrix} \mathbf{Q} & \mathbf{0} \\ \mathbf{0} & \mathbf{G} \end{pmatrix}\right)$. Premultiplied by matrices $\mathbf{R}_t$ and $\mathbf{N}_t$, these orthogonal disturbances transform into non-orthogonal residuals within each vector equation. The coefficients matrices $\mathbf{F}_t$, $\mathbf{C}_t'$, $\mathbf{R}_t$, $\mathbf{A}_t'$, $\mathbf{H}_t'$, $\mathbf{N}_t$, and the two variance-covariance matrices $\mathbf{Q}$ and $\mathbf{G}$ are estimated by maximizing the log-likelihood function of this system.

[6] Very detailed descriptions of the Kalman filter technique can be found in the Handbook of Econometrics by Hamilton [1994] and in his textbook [1994a]. Other useful contributions can be found in Aoki [1991], Harvey [1989], and Lütkepohl [1993].

3.2 Interpolation Models

3.2.1 Overview

In this section, we adapt the general state-space representation[7] (1) and (2) to our problem in different ways, specifically the inclusion of related series and assumed stochastic processes for monthly GDP. The interpolation framework[8] for $t = 1, \dots, T$ is:

$$\boldsymbol{\xi}_{t+1} = \mathbf{F}\boldsymbol{\xi}_t + \mathbf{C}'\mathbf{x}_{t+1} + \mathbf{R}\mathbf{v}_{t+1} \quad (3)$$
$$y_t^+ = \mathbf{a}_t'\mathbf{x}_t^* + \mathbf{h}_t' \cdot \boldsymbol{\xi}_t \quad (4)$$

On one hand, the state vector equation (3) describes the vector dynamics of the unobserved variable, monthly GDP (y_t), stacked in the state vector $\boldsymbol{\xi}_t = \begin{pmatrix} y_t & y_{t-1} & y_{t-2} \end{pmatrix}'$. The $[3 \times 1]$ dimension serves to take the three months within a quarter together in order to satisfy the sum-up constraint. The exact formulation of this vector equation is difficult, because there is no prior knowledge about the true process driving monthly GDP. In order to shed light on this issue, we compare various assumptions in section 5. We assume time-invariant coefficients for the matrices $\mathbf{F}$, $\mathbf{C}'$, and $\mathbf{R}$.

On the other hand, equation (4) relates the state vector to the observed quarterly GDP (y_t^+). Following Harvey and Pierse [1984], this observation equation represents the constraint that the sum of three monthly observations within a quarter must equal the quarterly observed GDP. Hence, this equality constraint implies that the error term disappears ($\mathbf{N}_t\mathbf{w}_t$) from the observation equation. The sum-up constraint is introduced by the coefficients vector $\mathbf{a}_t'$ and $\mathbf{h}_t'$ depending on the models in the following subsection[9].

All the specifications of the state-space models described hereafter correspond to different assumptions depending on the characteristics of the data to interpolate (driving process and stationarity) and whether related series are used or not. The properties of the data such as the order of integration and the stochastic process of monthly GDP influence the representation of the state equation. Possible related series influence the setup of the state vector equation and the

[7] The Kalman filter algorithm and the derivation of the log-likelihood function are displayed in Appendix A.

[8] In all the models, quarterly GDP $\left(y_t^+\right)$ is given each month, $y_1^+ = 0$, $y_2^+ = 0$, $y_3^+ =$ first quarterly value, $y_4^+ = 0$, $y_5^+ = 0$, $y_6^+ =$ second quarterly value, $y_7^+ = 0, \dots$, etc. Note that with T months to interpolate we observe $\frac{T}{3}$ quarterly values. But contrary to the convention when we stack the quarterly observations in one column vector to get $\mathbf{y}^+$, we do not include zero observations resulting thus in a vector of size $\left[\frac{T}{3} \times 1\right]$.

[9] Another alternative not treated in this text to introduce the sum-up constraint is to augment the state space representation with a "cumulator function" (y_t^c), which accumulates monthly GDP observations in a given quarter: $y_t^c = \sum_{s=0}^{r} y_{t-s}$ where $r = 0$ for $t = 1, 4, 7, \dots, T-2$, $r = 1$ for $t = 2, 5, 8, \dots, T-1$, and $r = 2$ for $t = 3, 6, 9, \dots, T$. See Harvey [15] for more details.

observation equation to in turn affect the coefficients contained in $\mathbf{C}'$ and $\mathbf{a}'_t$. We add related series in order to evaluate their statistical relevance. The selected assumptions are also guided by simplicity and technical considerations of the construction of the Kalman filter.

Hence, we focus on two broad classes of Kalman filter models summarized in figure 1.

Fig. 1 here

We classify our models with respect to the following questions. Do the models use related series ($\mathbf{x}_t$) or ($\mathbf{x}_t^*$) as explanatory variables? Does monthly GDP (y_t) follow an autoregressive (AR) process? If the data is not stationary, is there a correction for it? What does the structure of the monthly residuals look like? And finally, what is the algorithm for generating monthly GDP estimates?

The first class of models is designed without related series. We assume that there is enough information in the quarterly series, its autocovariance function, and in the assumed low order AR processes to generate monthly GDP. Moreover, we combine this assumption with alternative ways to treat non-stationary series (models **1a-c**). Contrasting to these AR models are "naive" models that neither follow an AR process nor include related series. However, it is not necessary to run the Kalman filter, because simple calculus produces the same results. For each quarter, model **1d** gives three equal monthly values, namely the quarterly mean. This indicates that the Kalman filter corrects the mean at each quarter taking into account a possible trend in the data. Model **1e** gives for each quarter three monthly GDP following a quarterly linear trend centered around the quarterly mean[10]. We take model **1e** as our benchmark because it has the simplest setup.

The second class of models introduces related series in order to extract information for the interpolation of monthly GDP. Within this group, we distinguish whether monthly GDP is allowed to follow an autoregressive process (models **2f-g**) or not (models **2a-e**). We further enrich this second class of models with different ways to treat non-stationarity and with different assumptions about monthly residuals.

In the next paragraphs we show the various models **1a–c** and **2a-g** in detail.

[10]Model **1d** needs a constant term as explanatory variable to calculate the quarterly mean. Model **1e** interpolates monthly observations linearly within a quarter, where we assume that we can split each quarter (except the first one) into an initial value y_{t-3} which is the last month of the previous quarter and a step d_t for $t = 4, 5, \ldots, T$ according to the following equation:

$$(\varphi y_{t-3} + d_t) + (\varphi y_{t-3} + 2d_t) + (\varphi y_{t-3} + 3d_t) = y_t^+$$

As the quarterly GDP (y_t^+), the step d_t is given each month, $d_4 = 0$, $d_5 = 0$, and $d_6 =$ second quarter step, etc. φ is a scalar that takes on 1 for $t = 6, 9, \ldots, T$ and 0 for $t = 4, 5, 7, \ldots, T-1$.

3.2.2 Models without Related Series

Model 1a In our first model, we assume that the first difference of monthly GDP follows a stationary AR(1) process: $\Delta y_t = \phi \Delta y_{t-1} + u_t$. Δy_t is the first difference of monthly GDP, ϕ is a coefficient constrained to lie inside the unit circle and u_t a iid error term with distribution $N(0, \sigma^2)$. In the treatment of the non-stationarity, we rewrite this AR(1) as an AR(2) of the series in levels:

$$y_t = (1+\phi)\, y_{t-1} - \phi y_{t-2} + u_t \tag{5}$$

This equation written in companion form, where $\boldsymbol{\xi}_t = \begin{pmatrix} y_t & y_{t-1} & y_{t-2} \end{pmatrix}'$, yields the state equation (6) for $t = 1, \ldots, T$.

$$\begin{pmatrix} y_{t+1} \\ y_t \\ y_{t-1} \end{pmatrix} = \begin{pmatrix} 1+\phi & -\phi & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} y_t \\ y_{t-1} \\ y_{t-2} \end{pmatrix} + \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} u_{t+1} \\ u_t \\ u_{t-1} \end{pmatrix} \tag{6}$$

Note that this formulation simply sets $\mathbf{C}' = \mathbf{0}$ in equation (3). The observation equation is simply incorporating the sum-up constraint leaving out related series. This implies $\mathbf{a}_t' = \mathbf{0}$ and $\mathbf{h}_t'$ taking on two different values depending on the respective month:

$$\begin{aligned} y_t^+ &= \mathbf{h}_t' \cdot \boldsymbol{\xi}_t \\ \text{where } \mathbf{h}_t' &= \begin{pmatrix} 0 & 0 & 0 \end{pmatrix}, \text{ for } t = 1, 2, 4, 5, 7, \ldots, T-1, \\ \text{where } \mathbf{h}_t' &= \begin{pmatrix} 1 & 1 & 1 \end{pmatrix}, \text{ for } t = 3, 6, 9, \ldots, T. \end{aligned} \tag{7}$$

Model 1b A recent alternative interpolative method was suggested by Bernanke, Gertler, and Watson [1997]. It consists in using GDP integrated of order one (I(1)) with a cointegrated series (p_t) such that we can compute a new monthly stationary series $y_t^s = \frac{y_t}{p_t}$ [11]. In fact p_t is just a scaling variable such that y_t^s is nontrending. One evident assumption underlying this approach is that the calculated multiplicative cointegration at a quarterly frequency will hold at a monthly frequency. For the dynamic specification of y_t^s, now forming the elements of state vector $\boldsymbol{\xi}_t$, we assume that it follows an AR(1) process $y_t^s = \phi y_{t-1}^s + u_t$. This results in a slightly different state equation than in model **1a** where y is replaced by y^s and the matrix $\mathbf{F}$ has a first row $\begin{pmatrix} \phi & 0 & 0 \end{pmatrix}$. The observation equation is also different because we have to "neutralize" the division by the I(1) series p_t. This is done by constraining the measurement equation in such a way that quarterly values of y_t^+ are restored in setting $\mathbf{a}_t' = \mathbf{0}$ and redefining vector $\mathbf{h}_t'$:

$$\begin{aligned} y_t^+ &= \mathbf{h}_t' \cdot \boldsymbol{\xi}_t \\ \text{where } \mathbf{h}_t' &= \begin{pmatrix} 0 & 0 & 0 \end{pmatrix}, \text{ for } t = 1, 2, 4, 5, 7, \ldots, T-1, \\ \text{where } \mathbf{h}_t' &= \begin{pmatrix} p_t & p_{t-1} & p_{t-2} \end{pmatrix}, \text{ for } t = 3, 6, 9, \ldots, T. \end{aligned} \tag{8}$$

[11] The ratio $y_t^s = \frac{y_t}{p_t}$ is chosen as a general framework in that it allows for negative values of the co-trending p_t.

Model 1c Bomhoff [1994] suggests to use the series in levels ignoring the stationarity problem, arguing that the Kalman filter does not require the user to make a definite decision regarding the need for differencing the data. The Kalman filter offers automatic processing capacity for a wide range of nonstationary time series. The models without related series and without AR processes have already suggested this feature. Hence, we could write down the law governing the process as if the series were stationary: $y_t = \phi y_{t-1} + u_t$. This model is similar to model **1a** but with a first row of matrix $\mathbf{F}$ defined as $\begin{pmatrix} \phi & 0 & 0 \end{pmatrix}$.

One main criticism of the models **1a-c** is that they extract signals only from the presumed stochastic process of the original series in a way that no new information is added. We could speak about "fool-yourself"-models[12] to generate the monthly GDP. It seems that we would be better off enriching the model with additional information. For this purpose, we now include new series that are related to the series to interpolate.

3.2.3 Models with Related Series and without AR Structure

Depending on the models, we introduce the related series either in the state vector equation (3) for the generalized least squares (GLS) estimator (models **2a-d**), or in the measurement equation (4) for the Kalman filter algorithm (models **2e-g**).

Model 2a-b Chow and Lin [1971] and [1976] showed how related series can be used to help interpolate lower frequency data in order to get higher frequency data with a GLS estimator. They assume that monthly GDP (y_t) can be described by a simple regression of y_t on w related series x_t, in matrix notation $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u}$, where the variance-covariance of the error term is $\mathbf{V} = E[\mathbf{u}\mathbf{u}']$. They also assume the same relationship at a quarterly level: $\mathbf{y}^+ = \mathbf{X}^+\boldsymbol{\beta} + \mathbf{u}^+$, where $\mathbf{X}^+$ is a matrix with quarterly average of three months of related series and $\mathbf{V}^+$ the variance-covariance matrix $E[\mathbf{u}^+\mathbf{u}^{+\prime}]$. $\mathbf{V}^+$ is a function of $\mathbf{V}$. We can configure the Kalman filter to match the Chow and Lin results defining vector $\boldsymbol{\xi}_t$ as suggested by Harvey and Pierse [1984]:

$$\begin{aligned} \boldsymbol{\xi}_{t+1} &= \mathbf{F}\boldsymbol{\xi}_t + \mathbf{R}\mathbf{v}_{t+1} && (9) \\ y_t^+ &= \mathbf{a}_t'\mathbf{x}_t^* + \mathbf{h}_t'\boldsymbol{\xi}_t && (10) \\ \text{where } \boldsymbol{\xi}_t &= \begin{pmatrix} y_t - \mathbf{x}_t'\boldsymbol{\beta} \\ y_{t-1} - \mathbf{x}_{t-1}'\boldsymbol{\beta} \\ y_{t-2} - \mathbf{x}_{t-2}'\boldsymbol{\beta} \end{pmatrix}, \mathbf{F} = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \mathbf{x}_t^* = \sum_{j=t-2}^{t} \mathbf{x}_j, \\ \mathbf{h}_t' &= \begin{pmatrix} 0 & 0 & 0 \end{pmatrix} \text{ and } \mathbf{a}_t' = \mathbf{0} \text{ for } t = 1, 2, 4, 5, 7, \ldots, T-1, \\ \mathbf{h}_t' &= \begin{pmatrix} 1 & 1 & 1 \end{pmatrix} \text{ and } \mathbf{a}_t' = \boldsymbol{\beta}' \text{ for } t = 3, 6, 9, \ldots, T. \end{aligned}$$

$\mathbf{R}\mathbf{v}_t$ is equal to $\begin{pmatrix} u_t & 0 & 0 \end{pmatrix}'$. As the related series ($\mathbf{x}_t$) are implicitly in the state vector equation, we set $\mathbf{C}' = \mathbf{0}$ and reintroduce them as $\mathbf{x}_t^*$ in the observation

[12] Thanks to Mark Watson for bringing up this expression.

equation. In their seminal paper, Chow and Lin directly calculate this best linear unbiased estimator $\hat{y}_t$ for the monthly series from the trace minimization of the covariance matrix $Cov(\hat{\mathbf{y}} - \mathbf{y})$. They do not use the Kalman iterations and so avoid the problems associated to the numerical optimization procedures maximizing the log-likelihood function. This Kalman filter configuration and the following Chow and Lin regression yield the same log-likelihood function[13]. The estimates of monthly GDP $\hat{y}_t$ are:

$$\hat{\mathbf{y}} = \mathbf{X}\hat{\boldsymbol{\beta}}_{GLS} + \boldsymbol{\Lambda}(\mathbf{V})\hat{\mathbf{u}}^{+} \tag{11}$$

where $\hat{\mathbf{y}}$ denotes the vector of the monthly variables in matrix notation. This special fitted value has two parts: a traditional fitted value $\mathbf{X}\hat{\boldsymbol{\beta}}_{GLS}$ with the influence of related series and an interpolation-corrected residuals term $\boldsymbol{\Lambda}(\mathbf{V})\hat{\mathbf{u}}^{+}$. $\hat{\boldsymbol{\beta}}_{GLS}$ is a GLS estimator corresponding to the regression between the quarterly GDP data ($\mathbf{y}^{+}$) and the "quarterly related series" ($\mathbf{X}^{+}$):

$$\hat{\boldsymbol{\beta}}_{GLS} = \left(\mathbf{X}^{+\prime}\mathbf{V}^{+^{-1}}\mathbf{X}^{+}\right)^{-1}\mathbf{X}^{+\prime}\mathbf{V}^{+^{-1}}\mathbf{y}^{+} \tag{12}$$

The weighting matrix in the GLS regression is the inverse of the variance-covariance matrix $\mathbf{V}^{+}$ of the quarterly residuals $\mathbf{u}^{+}$. Hence, the assumptions about $\mathbf{V}$ directly influence the distribution of $\hat{\boldsymbol{\beta}}_{GLS}$ and the $\left[T \times \frac{T}{3}\right]$ matrix $\boldsymbol{\Lambda}$ for the dissemination of the quarterly residuals over the monthly estimated GDP. These quarterly residuals are crucial for interpolation, because the traditional fitted monthly values $\mathbf{X}\hat{\boldsymbol{\beta}}_{GLS}$ do not sum up to quarterly observations. Therefore, the residuals $\mathbf{u}^{+}$ must be "redistributed" along the monthly GDP according to the weighting matrix $\boldsymbol{\Lambda}$ to correct this shortcoming.

Models **2a** and **2b** differ in $\mathbf{V}$. In model **2a,** we assume that the variance-covariance of the monthly residuals $\mathbf{V} = E\left[\mathbf{u}\mathbf{u}'\right]$ is a simple diagonal matrix $\sigma_u^2\mathbf{I}_T$. It implies that $\mathbf{V}^{+}$ the variance-covariance matrix of quarterly residuals is equal to $E\left[\mathbf{u}^{+}\mathbf{u}^{+\prime}\right] = \frac{\sigma_u^2}{3}\mathbf{I}_{\frac{T}{3}}$ and $\boldsymbol{\Lambda}$ is equal to

$$\begin{pmatrix} 1 & 0 & \dots & & \\ 1 & 0 & & & \\ 1 & 0 & & & \vdots \\ \vdots & & & 0 & 1 \\ & & & 0 & 1 \\ & & \dots & 0 & 1 \end{pmatrix}. \tag{13}$$

However, the diagonal variance-covariance matrix $\mathbf{V} = E\left[\mathbf{u}\mathbf{u}'\right]$ is rarely supported by the data. A way to improve this setting is to allow for serial correlation

[13] Appendix B describes the Chow and Lin regression and Appendix C shows that the Kalman filter and the Chow and Lin regression yield the same estimates by maximum likelihood.

in the error term. Hence, we assume for the model **2b** that the error term follows an AR(1), $u_t = \varrho u_{t-1} + \varepsilon_t$ where ε_t is a white noise, yielding a variance-covariance matrix $\mathbf{V}$:

$$\begin{pmatrix} 1 & \varrho & \varrho^2 & \dots & \varrho^{T-1} \\ \varrho & 1 & \varrho & & \\ \varrho^2 & \varrho & 1 & & \\ \vdots & & & & \vdots \\ \varrho^{T-1} & & & \dots & 1 \end{pmatrix} \frac{\sigma_\varepsilon^2}{1-\varrho^2}. \tag{14}$$

This specification introduces a different $\hat{\boldsymbol{\beta}}_{GLS}$ and a new redistribution matrix $\boldsymbol{\Lambda}$ dependent on ϱ. The more ϱ tends to zero, the more the $\boldsymbol{\Lambda}$ matrix converges towards matrix (13). Hence, if the autocorrelation is significant, redistribution is less "rigid" than in model **2a** and the quarterly residuals are not only spread out over their corresponding months but also influence monthly GDP of surrounding quarters in a "smoother" way.

Model 2c-d A variation of models **2a-b**, as suggested by Denton [1971] and Fernandez [1981], is to use first differenced time series in the regression instead of levels in order to account for non-stationarity. They assume that the variance-covariance of the error term is $\mathbf{V} = E[\mathbf{uu}'] = \sigma^2 \mathbf{I}_T$ in model **2c** or that the error term follows an AR(1) yielding a variance-covariance matrix $\mathbf{V}$ equal to the matrix (14) in model **2d** . They compute $\hat{\boldsymbol{\beta}}_{GLS}$ with a weighting matrix equal to the inverse of the quarterly equivalent of $(\mathbf{D}'\mathbf{V}'\mathbf{D})^{-1}$ where:

$$\mathbf{D} = \begin{pmatrix} 1 & 0 & 0 & \dots \\ -1 & 1 & 0 & \\ 0 & -1 & 1 & \\ 0 & 0 & -1 & \\ \vdots & & & \ddots \end{pmatrix}, (\mathbf{D}'\mathbf{V}'\mathbf{D})^{-1} = \begin{pmatrix} 1 & 1 & 1 & 1 & \dots & 1 \\ 1 & 2 & 2 & 2 & & \\ 1 & 2 & 3 & 3 & & \\ 1 & 2 & 3 & 4 & & \\ \vdots & & & & \ddots & \\ 1 & & & & & T \end{pmatrix}. \tag{15}$$

Matrix $\mathbf{D}$ is a first difference operator. From this follows that the $\boldsymbol{\Lambda}$ matrix implies for both models a new redistribution of the quarterly residuals, where weighted moving average of quarterly residuals is given to each monthly estimated GDP.

Model 2e This model uses the Kalman filter algorithm. The advantage is that we do not have to assume the driving process for the monthly residuals. The Kalman filter considers the most appropriate one. As in the Chow and Lin regression, we assume no autoregressive process to see the improvement of the Kalman Filter and its greater flexibility in shaping its equivalent of $\boldsymbol{\Lambda}(\mathbf{V})\hat{\mathbf{u}}^+$.

The state-space form is the following:

$$
\begin{aligned}
\begin{pmatrix} y_{t+1} \\ y_t \\ y_{t-1} \end{pmatrix} &= \begin{pmatrix} c_1 & c_2 & \dots & c_w \\ 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \end{pmatrix} \begin{pmatrix} x^1_{t+1} \\ x^2_{t+1} \\ \vdots \\ x^w_{t+1} \end{pmatrix} \\
&+ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} u_{t+1} \\ u_t \\ u_{t-1} \end{pmatrix} \qquad (16) \\
y_t^+ &= \mathbf{h}_t' \cdot \boldsymbol{\xi}_t \qquad (17) \\
\text{where } \mathbf{h}_t' &= \begin{pmatrix} 0 & 0 & 0 \end{pmatrix}, \text{ for } t = 1, 2, 4, 5, 7, \dots, T-1, \\
\text{where } \mathbf{h}_t' &= \begin{pmatrix} 1 & 1 & 1 \end{pmatrix}, \text{ for } t = 3, 6, 9, \dots, T.
\end{aligned}
$$

In equation (16) the autoregressive part disappears from the state equation and the related series are introduced by the $[3 \times w]$ matrix $\mathbf{C}'$. The error terms for the state equation are standard. We estimate the c coefficients by numerically maximizing the log-likelihood function.

3.2.4 Models with Related Series and with AR Structure

Model 2f In addition to the related series, we assume here that the autoregressive structure explained in the previous class of models helps to describe the monthly GDP. The non-stationarity correction is similar to the one done in the model **1a** with model in AR(2) form. After the inclusion of related series to model **1a**, the state equation becomes equation (6) plus the element $\mathbf{C}'\mathbf{x}_{t+1}$ where $\mathbf{x}_{t+1}$ includes the w related series.

Model 2g This model is similar to model **1b** but with the addition of related series as in model **2f** with a corresponding matrix $\mathbf{C}'$.

4 Data

4.1 Signal Extraction from Related Series

A key factor in the present interpolation problem is the signal extraction from related series. Besides the assumption about the dynamics of GDP, related time series data represent the main information source for interpolation. These data must fulfill two requirements.

First, they need to be correlated with the series to interpolate. The higher the systematic comovements with GDP are, the stronger is the signal that can be exploited to fill the gaps. On the other hand, if there is only a modest information content in the related series, this comes at the cost of a lot of noise that is introduced in the interpolated series. The choice of the related series is therefore crucial in order to successfully estimate a series at higher frequency.

Second, the related series need to be available in the desired higher frequency of the interpolated GDP. The fact that there are not many macroeconomic series available at monthly frequency imposes a strong restriction in Switzerland. This leads us to use other than Swiss variables that we assume to be highly correlated with the desired related series.

These two points require a thorough investigation for the task of choosing the correct related series. Amemiya [1980] suggests a joint strategy based on economic-theoretic considerations and on statistical evidence. Economic intuition can often indicate which data series to choose and what functional form they should have. Together with this, it is convenient to have a single statistical measure to choose related series that produce the "best" result. These two aspects, intuitive approach and choice metrics, should be viewed as forming a single choice package rather than being in competition with each other. They allow to make a final choice of the series which we use in our models. Both elements of the selection process will be presented in detail in the following section.

4.2 Choice of Related Series

4.2.1 Economic Intuition

The most natural way to approach the series selection problem is to split up GDP into its expenditure components, private consumption (C), private domestic investments (I), government expenses (G) and net exports ($X - M$):

$$Y = C + I + G + X - M \tag{18}$$

With the exception of exports and imports, none of these series is available at the higher frequency. Therefore, it is necessary to identify related data series that proxy for the desired components.

An alternative to breaking GDP into its expenditure components is to benefit from the characteristics of Switzerland as a small open economy and the important comovement between domestic and foreign business cycles. Taking into consideration monthly foreign main economic indicators allows us to choose the related series from a broader data set as Switzerland's closest trade partners have traditionally large statistical databases.

The existence of two groups of data series, national accounting and open economy, imply that the pure evaluation of potential candidates within each group has to be complemented with the comparison of entire sets of related series after the different model estimations[14].

[14]From the limited degrees of freedom in applying economic theories due to data availability restrictions, it follows that the statistical evaluation must take a more important place than it usually would according to Amemiya [1980].

4.2.2 Statistical Evaluation

The case discussed here is the search for individual proxy variables in economic models[15]. Suppose, you identify a set of related data series $\mathbf{x}$ out of which variable x_k is unobservable. Furthermore, the variable y which is being interpolated depends linearly on $\mathbf{x}$.

$$y_t = \alpha_0 + \alpha_1 x_{1,t} + \alpha_2 x_{2,t} + \ldots + \alpha_k x_{k,t} + \ldots + u_t \tag{19}$$

The goal is to choose the best observable proxy for x_k. In cases like this, an informal method often applied is replacing x_k with the variable z_k which yields the highest R^2 of all possible variables z in equation (19). Leamer [1983] shows that if the proxy variables z are assumed to depend linearly on x_k and the error terms being Niid, the best proxy is the one that produces the highest R^2. In the univariate regression $z_{i,t} = \delta_i x_{k,t} + \varepsilon_{i,t}$, the particular z_i which yields the smallest variance $\sigma^2_{\varepsilon_i}$ could be defined as the best proxy. Leamer [1983] uses a likelihood ratio test to show the unambiguously negative relationship between the variance of the error term and the R^2.

Another popular method which can be applied to a wider range of competing models than the one R^2 criterion above is the method of penalized likelihood. The best known examples in this class of criteria which has grown a lot in the last twenty years are the Akaike Information Criterion (AIC) [1974] and the Schwarz Information Criterion (SIC) [1978]. In this class of criteria, a term that acts to punish additional coefficients is added to the likelihood function.

4.3 Data Description

For a long time, Switzerland has stayed far behind other European countries in the development of economic statistical data. In 1996, as part of a reform program, national accounting was adapted to the European System of National Accounting (ESNA) 78[16]. Thereafter, GDP was calculated differently. The Federal Statistics Office dated the series back to 1980 such that there is now a data sample of more than 18 years or 73 quarterly observations. The figures to be interpolated are deflated and deseasonalized[17].

The related series[18] in the national accounting approach have been identified as retail sales (RS) to proxy for private consumption and as the level of not

[15] At this stage of the text, we describe the data selection within one group of related series as described in section 4.2.1. Choosing the best set of related series is also independent of the different models presented in section 3.

[16] Up to and including 1996, Swiss GDP was recorded following the OECD standard 58. According to Federal Statistics Office, it is planned that to adopt the ESNA 95 standard within a few years.

[17] Deseasonalization was executed using the X12/ARIMA method of the US Bureau of Census.

[18] All the series, with the exception of real GDP given by the Federal Office for Economic Development and Labour, are provided by Datastream.

utilized construction loans to inversely proxy for investment (NL). These monthly available proxies have been selected based on the three criteria described in the previous section. Furthermore, we include exports (X) and imports (M). All the series are entered in levels[19]. Government expenditure was dropped in the national accounting approach due to its low covariance with the business cycle. This would have introduced too much noise and moreover, there is no sensible proxy at monthly frequency for it.

As foreign series, we use a composite index of industrial production (COMIP)[20], British IP (UKIP), and German IP (BRDIP). IP are the foreign monthly available series that move closest with the Swiss business cycle of all the related foreign series considered (results not reported).

Prior to estimation, we have excluded several potential series based on economic arguments or on the statistical evaluation of the previous section. French IP, Italian IP, survey data by the KOF[21], labor market figures, exchange rates and commodity prices were eliminated statistically. We have neither included variables that have proved to have predicting power for GDP such as the term spread because of unrealistic assumptions on the lead-lag relationship that would have been necessary. Figure 2 and table 1 give an overlook over the series used in this paper.

Fig. 2 and Table 1 here

During the 18 years of observations, the state of the Swiss economy can be roughly divided in two parts. Figure 2 clearly shows the phases of economic growth and prosperity in the 1980's and of stagnation in the 1990's. During its recession, Switzerland exhibited the lowest real GDP growth of all European countries[22].

Table 1 reports basic summary statistics of the quarterly and monthly series that will be used for interpolation. Following the integration results from figure 2 and from augmented Dickey-Fuller (ADF) tests for all the variables (not reported), we find that the levels of all the series are non-stationary. Hence, we report the results for growth rates. The ADF tests and the AR(1) regressions on the growth rates confirm that the levels of the series are not stationary. The different values of the contemporary cross-correlations also confirm the requirement of the comovements of the related series with the quarterly GDP. Finally, these cross-correlations also show why we only consider contemporary relationships between the related series and the quarterly GDP. It is actually very difficult to find

[19]The models transform the level vectors into the desired form.

[20]IP of five countries (major trade partners of Switzerland) are weighted according to the share of Swiss exports to the respective countries in 1996.

[21]Institute for Business Cycle Research of the Swiss Federal Institute of Technology.

[22]To keep things simple and for further research, we decided not to take into account this structural break that would mean to combine our interpolations models with time-varying parameters generally dealt within the Kalman filter framework or with the use of dummy variables.

robust leads and lags - the so-called stylized facts of the business cycles literature - between GDP and our proxy variables.

We also perform a Johansen [1991] test to check for cointegration that is needed for the evaluation of the applicability of the Bernanke, Gertler, and Watson extension [1997]. It is natural to assume that RS is moving along with GDP. We therefore test the quarterly proxy for cointegration. The test results reject the hypothesis of no cointegration at the 1% significance level. These results are reported in table 2.

Table 2 here

Not reported are the tests of other potentially cointegrated variables with an economic interpretation. All the tests reveal that only the quarterly GDP and the volume of retail sales (RS) are cointegrated. Hence, we use RS either as a related series or as the detrending series (p_t) in the Bernanke, Gertler, and Watson framework[23]. In fact we cannot directly test the needed multiplicative cointegration, but a ADF test of the stationarity of the quarterly equivalent of $y_t^s = \frac{y_t}{p_t} = \frac{GDP_t}{RS_t}$ reveals that this ratio is stationary at the 1% significance level.

5 Results

5.1 Overview

The interpolation results are displayed in table 3. It contains for each model statistical information about the estimated series for the period 1981-1997[24], the related series, the information criterion, the log-likelihood, and key indicators for the annualized growth rate of the monthly interpolated GDP. Two mean square errors (MSE) for the evaluation of the models are given. The first one is between the level of the interpolated benchmark (model **1e**) and the interpolated series of each model, respectively. The second one is the MSE between the true quarterly GDP and a simulated quarterly interpolated GDP from annual data within the model in question in order to compare how the interpolation model would have performed at a frequency where models can be selected unambiguously based on an available data set.

Table 3 here

Note, that table 3 is constructed in order to evaluate the models with respect to two basic directions. First, it is important to know if the inclusion of related

[23]To prevent the detrending series from introducing excessive volatility in the system we take only the low frequency part of RS after Hodrick-Prescott filtering. The main objective of detrending GDP can still be maintained.

[24]Due to initial oscillations we discard twelve months of observations which otherwise would have heavily influenced the results.

series (class **2**) performs better than the "fool-yourself" class **1**. Second, we investigate the appropriate treatment of non-stationarity and analyze the question whether applying modern techniques perform better than traditional ones.

5.2 Evaluation of Related Series

Our point of view is that it is desirable to have an economic model underlying the interpolation instead of just a purely econometric procedure which appears very mechanical. Moreover, related series could possibly break the regular pattern within a quarter produced by all interpolation procedures without related series[25] as shown in figure 3.

Fig. 3 here

In its left hand part, the figure shows the plot of series **20** estimates and the published quarterly GDP estimates[26]. The cyclical pattern within the quarter is illustrated on the right hand side as an average difference for the three months within the quarter between series **20** and the benchmark for growth and decline periods respectively. The deviations are significant for the first and the last observation within the quarter and leads us to reject the model for economic reasons. The presence of the pattern in the GDP estimates is due to the autocorrelation coefficient and its shape depends on the sign of GDP growth. It is thus tempting to include related series to eliminate the pattern produced by the autoregressive structure of the model. However, we found that in all type **2f** series the inclusion of related series, relative to model **1a**, does not attenuate the pattern but exacerbate it. Thus, the only way to eliminate the pattern is to remove its source, the autoregressive structure, and to use models **2a-e** which assume no autocorrelation. In these models we can see the related series implementing their movements in the interpolated GDP leading to the desired oscillations.

Econometrically, the conclusion of whether to include related series or not is ambiguous. For series **15** to **19**, the value of the AIC is higher than for their respective base case, **1d** (not reported) and **1a,** while only for series **20** it is lower. Generally, testing the significance using a likelihood ratio test shows that introducing related series does improve the performance of the interpolation for the model **2e** but not for model **2f** where the related series are just entered as growth rates. Results show that related series always come at a cost of introducing noise in the interpolated series. In our case, a lot of monthly series display too much noise in order to make an economic sense. The standard deviation of

[25] The pattern is systematically convex or concave if the model has an autoregressive structure, depending on growth state of the economy. Monthly GDP estimates produced by model **1e** are linear and model **1d** produces monthly estimates which equal 1/3 of quarterly GDP.

[26] The three months of each quarter sum to the value of this quarter. However, the line of the monthly interpolated GDP does not exactly pass through the points of quarterly GDP as the latter is simply scaled by a factor of 1/3. The fact that the dots are not exactly on the line cannot be interpreted as a quality indicator.

the growth rates of different interpolated series is a good indicator. Hence, all the models generating too much volatility relative to the annualized standard deviation of the quarterly GDP estimates are not displayed[27] in table 3.

Regarding the two sets of related series, one observes that in general related series based on the open economy assumption introduce less volatility in the generated growth rates than the national accounting variables. For the reported related series COMIP[28] and NL this relation is reversed. However, including additional variables in the national accounting approach increases the volatility considerably. To further investigate the characteristics of the most appropriate related series note that within each model the log-likelihood values show that the national accounting approach is preferable even if not always significantly.

Another evaluation criteria is the mean squared error (MSE) of a model series with respect to the benchmark. The results indicate that in general adding related series increases the MSE reflecting an increase in volatility as the models deviate more from the smooth benchmark. As this criteria is a rather soft one and as there are models with the contrary effect, it does not seem suitable for model evaluation. Moreover, our benchmark is only founded on practical reasons and hence, cannot be regarded as an objective measure for model evaluation.

5.3 Evaluation of Techniques

The comparison between different interpolation setups and the question whether modern methods perform better than traditional ones is closely linked to the treatment of stationarity. First of all, within the regression based methods the correction for nonstationarity proposed by Denton and Fernandez (DF, model **2c**) does produce results that are qualitatively only slightly better than the classic Chow and Lin method using level series (CL, model **2a**). The effect of modelling AR(1) error terms in the CL-model (model **2b**) and in the DF-model (model **2d**) is not clear. In the CL-models, the likelihood falls while for the DF-models it increases, when AR(1) error terms are considered. The standard deviation of the generated series rises in the CL-models and behaves irregularly in the DF-setups. This shows that these methods achieve a higher likelihood when series are more volatile.

The theoretical advantages of the Kalman Filter concerning the assumptions about the treatment of nonstationarity, the error terms of the process govern-

[27] We restrict ourselves to models that produce series with an annualized standard deviation lower than four times the variability of the growth rate of the official quarterly GDP estimates (13%). Comparisons between monthly and quarterly values of industrial production growth in various countries show that the annualized values of monthly standard deviation are 2 to 4 times higher than quarterly ones which serve as a reference.

[28] Of all the series that could not be distinguished by statistical evaluation in section 4.2.2, COMIP is found to be the most useful related series of the open economy approach. Results using BRDIP and UKIP are therefore not reported in table 3.

ing monthly GDP and the dynamic updating create high expectations for the interpolation results that cannot all be satisfactorily met when monthly GDP is analyzed in Switzerland.

Models constructed in a first difference equation with an autoregressive structure suggested by Bernanke, Gertler, and Watson [1997] are clearly worse than the ones reported in table 3, both in terms of cyclical regularity and in volatility. This procedure neglects the fact that the Kalman filter already corrects the nonstationarity of the data (Lütkepohl [1993]), creating a kind of redundance in the correction of this phenomena. The model **2e** is in this sense a better model. We do not explicitly account for nonstationarity but keep the improved treatment of the error terms with respect to the traditional models. Model **2e** is in fact similar to models **2a-b**, but with a wider range of error terms. The filter picks up the best model of error terms within the full range of possible constructs, without the need to state their characteristics.

5.4 A Monthly GDP Estimate

Based on this mixed evidence concerning the two directions, we recommend the series **17** for further research. It has the advantage to be produced by a Kalman filter framework including related series. Due the absence of autoregressive structure, it does not display a regular pattern and the series does exhibit moderate volatility. The plot is given in figure 4 and the values in table 4 in the next section.

Can this extensive selection procedure be confirmed by first interpolating annual to quarterly data and then comparing the resulting quarterly series with the true GDP estimates? If yes, then we would have a very handy tool for the evaluation of competing interpolation models. Of course, the underlying assumption that the best annual interpolation model is also the best quarterly one is strong, but if the criteria does well, it could well be used as suggestive evidence in similar problems. Moreover, there is no reason to think that the frequency change has a fundamental impact on the performance of the models[29]. But the results show that it is not always the case that models with highest likelihood are the best interpolating models at the lower frequency. Within the GLS based class just model **2c** confirms our expectations. For all models with a pattern, applying this method makes no sense. However, for our selected series the suggestive evidence is partly verified. Therefore, we conclude that this approach may be used as an indicator only but certainly not as a selection criteria.

[29] Another way to apply this proposal would be to select the model with the best AIC for the interpolation from annual data and to see if the same model also produces the best AIC for the interpolation of monthly data from quarterly data.

6 Conclusion

In this paper, we describe a setup that nests a wide range of interpolation models in the literature and we apply it to Swiss GDP. The goal of this paper is to evaluate alternative interpolation models and then to produce a monthly deseasonalized real GDP available for researchers and practitioners. The results are given in figure 4 and table 4.

Fig. 4 and Table 4 here

With respect to the nonstationarity and to the usefulness of related series, it is difficult to a priori present a clear-cut answer how these puzzles can be treated best. Our results show that the ways to consider the nonstationarity problem and to solve it with the second-order autoregressive structure (models **1a**, **2f**) or with the detrending method (models **1b**, **2g**) are not suitable for Swiss data. These two methods impose econometric characteristics on the produced data that cannot be carried further for an economic interpretation. The nonstationarity correction made by the filter itself seems to be sufficient.

Our results further show that in particular cases, related series can be very useful. In this case, economic interpretation backed on our two economic models is only based on the comparison between the volatilities of the growth rates of the quarterly values and of the computed monthly series, and some subsidiary indicators. However, including related series does not systematically improve the results of the base case.

The data does not seem to confirm unambiguously the expected long–run hypothesis between the interpolation at a monthly and at a quarterly level. A more rigorous econometric analysis would be needed to know if this comparison transgresses short-run considerations.

For the interpolation of Swiss GDP we suggest to use an approach with the four related series exports, imports, retail sales, and not utilized construction loans, the latter two of which are proxying for consumption and investment which are not monthly recorded. Furthermore, we find that the best results can be achieved in a Kalman filter framework with no restrictions on the error terms.

References

[1] Hirotugu Akaike. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, AC-19:716–723, 1974.

[2] Takeshi Amemiya. Selection of regressors. *International Economic Review*, 21(2):331–354, June 1980.

[3] Masano Aoki and Arthur Havenner. State space modeling of multiple time series. *Econometric Review*, 10(1):1–59, 1991.

[4] Ben S. Bernanke, Mark Gertler, and Mark Watson. Systemic monetary policy and the effects of oil price shocks. *Brookings Papers on Economic Activity*, (1):91–157, 1997.

[5] Eduard J. Bomhoff. *Financial Forecasting for Business and Economics*. Dryden, 1994.

[6] Gregory C. Chow and An-Loh Lin. Best linear unbiased interpolation, distribution, and extrapolation of time series by related series. *The Review of Economics and Statistics*, 53(4):372–375, 1971.

[7] Gregory C. Chow and An-Loh Lin. Best linear unbiased estimation of missing observation in an economic time series. *Journal of the American Statistical Association*, 71(355):719–721, September 1976.

[8] Enrique de Alba. Estimación del PIB trimestral para México. *Estudios Económicos*, 5(2):359–370, July-December 1990.

[9] Frank T. Denton. Adjustment of monthly or quarterly series to annual totals. *Journal of the American Statistical Association*, 66(333):99–102, 1971.

[10] Roque B. Fernandez. A methodological note on the estimation of time series. *The Review of Economics and Statistics*, 63(3):471–476, 1981.

[11] Milton Friedman. The interpolation of time series by related series. *Journal of the American Statistical Association*, 57(300):729–757, 1962.

[12] Richard Guay, Ross D. Milbourne, Glenn Otto, and Gregor W. Smith. Estimation du PIB mensuel canadien: 1962 à 1985. *L'Actualité économique, Revue d'analyse économique*, 66(1):14–30, 1990.

[13] James D. Hamilton. State-space models. Handbook of Econometrics IV ed. by R. F. Engle and D. L. McFadden, Elsevier, 1994.

[14] James D. Hamilton. *Time Series Analysis*. Princeton University Press, 1994a.

[15] Andrew C. Harvey. *Forecasting, Structural Series Models and the Kalman Filter*. Cambridge University Press, 1989.

[16] Andrew C. Harvey and Richard G. Pierse. Estimating missing observations in economic time series. *Journal of the American Statistical Association*, 79(385):125–131, March 1984.

[17] Soren Johansen. Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models. *Econometrica*, 59:1551–80, November 1991.

[18] Rudolf E. Kalman. A new approach to linear filtering and prediction problems. *Journal of Basic Engineering, Transactions of the ASME Series D*, 82:35–45, 1960.

[19] Rudolf E. Kalman. New methods in Wiener filtering theory. Proceedings of the first symposium of engineering applications of random function theory and probability ed. by John L. Bogdanoff and Frank Kozin, Wiley, 1963.

[20] Steven G. Lanning. Missing observations: A simultaneous approach versus interpolation by related series. *Journal of Economic and Social Measurement*, 14(1):155–163, April 1986.

[21] Edward E. Leamer. Model choice and specification analysis. Handbook of Econometrics I ed. by Z. Griliches and M. D. Intriligator, North-Holland, 1983.

[22] Robert B. Litterman. A random walk, Markov model for the distribution of time series. *Journal of Business and Economic Statistics*, 1(2):169–173, 1983.

[23] Helmut Lütkepohl. *Introduction to Multiple Time Series Analysis*. Springer, 2nd edition, 1993.

[24] Maximiano Pinheiro and Carlos Coimbra. Distribution and extrapolation of time series by related series using logarithms and smoothing penalties. *Economia*, 12(3):359–374, October 1993.

[25] James R. Schmidt. A general framework for interpolation, distribution, and extrapolation of a time series by related series. Regional econometric modeling ed. by M. Ray Perryman and James R. Schmidt, Kluwer, 1986.

[26] André Schwaller and Bruno Parnisari. Die Quartalsschätzungen des Bruttoinlandproduktes auf Grundlage der revidierten Volkswirtschaftlichen Gesamtrechnung (ESVG78). *Mitteilungsblatt für Konjunkturfragen, Federal Office for Economic Development and Labour*, (1):3–24, Fall 1997.

[27] Gideon Schwarz. Estimating the dimension of a model. *Annals of Statistics*, 6:147–164, 1978.

Appendix A: Kalman Filter

In this appendix, we show the iteration steps of the Kalman filter. We also give the log-likelihood function of our system. All our interpolation models, are based on equations (3) and (4), $\boldsymbol{\xi}_{t+1} = \mathbf{F}\boldsymbol{\xi}_t + \mathbf{C}'\mathbf{x}_{t+1} + \mathbf{R}\mathbf{v}_{t+1}$, and $y_t^+ = \mathbf{a}_t'\mathbf{x}_t^* + \mathbf{h}_t' \cdot \boldsymbol{\xi}_t$. The Kalman filter iteration, correction, prediction, and MSE steps at time t, is the following loop. At time t assume that $y_0^+, y_1^+ y_2^+, \ldots, y_{t-1}^+, y_t^+$ are known. The related series $\mathbf{x}$ and $\mathbf{x}^*$ are known but up to $t+1$. The predictions at time $t-1$ are also known: $\hat{\boldsymbol{\xi}}_{t|t-1}, \hat{y}_{t|t-1}^+$. The MSE are also known: $\mathbf{P}_{t|t-1} = \text{MSE}\left(\hat{\boldsymbol{\xi}}_{t|t-1}\right)$, and $\text{MSE}\left(\hat{y}_{t|t-1}^+\right)$.

Correction or update step:

- $\hat{\boldsymbol{\xi}}_{t|t} = \hat{\boldsymbol{\xi}}_{t|t-1} + \underbrace{\mathbf{P}_{t|t-1}\mathbf{h}_t \left(\text{MSE}\left(\hat{y}_{t|t-1}^+\right)\right)^{-1}}_{Gain} \cdot \left(y_t^+ - \hat{y}_{t|t-1}^+\right)$
- $\mathbf{P}_{t|t} = \mathbf{P}_{t|t-1} - \mathbf{P}_{t|t-1}\mathbf{h}_t \left(\text{MSE}\left(\hat{y}_{t|t-1}^+\right)\right)^{-1} \mathbf{h}_t'\mathbf{P}_{t|t-1}$

Prediction step:

- Prediction step: $\hat{\boldsymbol{\xi}}_{t+1|t} = \mathbf{F}\hat{\boldsymbol{\xi}}_{t|t} + \mathbf{C}' \cdot \mathbf{x}_t$
- Prediction step: $\hat{y}_{t+1|t}^+ = \mathbf{a}_{t+1}' \cdot \mathbf{x}_{t+1}^* + \mathbf{h}_{t+1}' \cdot \hat{\boldsymbol{\xi}}_{t+1|t}$

MSE step:

- $\text{MSE}\left(\hat{\boldsymbol{\xi}}_{t+1|t}\right) = \mathbf{P}_{t+1|t} = \mathbf{F} \cdot \mathbf{P}_{t|t} \cdot \mathbf{F}' + \mathbf{RQR}'$
- $\text{MSE}\left(\hat{y}_{t+1|t}^+\right) = \mathbf{h}_{t+1}'\mathbf{P}_{t+1|t}\mathbf{h}_{t+1}$

Log-likelihood function:

Each observation of the sample y_t^+ is normally distributed:
$y_t^+ \mid \left(y_0^+, y_1^+, y_2^+, \ldots, y_{t-1}^+, \mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_t, \mathbf{x}_1^*, \mathbf{x}_2^*, \ldots, \mathbf{x}_t^*\right)$
$\sim N\left(\left(\mathbf{a}_t' \cdot \mathbf{x}_t^* + \mathbf{h}_t' \cdot \hat{\boldsymbol{\xi}}_{t|t-1}\right), \left(\mathbf{h}_t'\mathbf{P}_{t|t-1}\mathbf{h}_t\right)\right)$

The log-likelihood function for the whole sample is the following expression:

$$\sum_{t=1}^{T} \ln f\left(y_t^+\right) =$$
$$-\frac{T}{2}\ln(2\pi) - \frac{1}{2}\sum_{t=1}^{T} \ln\left|\mathbf{h}_t'\mathbf{P}_{t|t-1}\mathbf{h}_t\right|$$
$$-\frac{1}{2}\sum_{t=1}^{T}\left[\begin{array}{c} \left(y_t^+ - \mathbf{a}_t' \cdot \mathbf{x}_t^* - \mathbf{h}_t' \cdot \hat{\boldsymbol{\xi}}_{t|t-1}\right)\left(\mathbf{h}_t'\mathbf{P}_{t|t-1}\mathbf{h}_t\right)^{-1} \cdot \\ \left(y_t^+ - \mathbf{a}_t' \cdot \mathbf{x}_t^* - \mathbf{h}_t' \cdot \hat{\boldsymbol{\xi}}_{t|t-1}\right)' \end{array}\right]$$

$$
\begin{aligned}
&\sum_{t=1}^{T} \ln f\left(y_t^{+}\right) = \\
&-\frac{T}{2}\ln\left(2\pi\right) - \frac{1}{2}\sum_{t=1}^{T} \ln\left|\mathbf{h}_t'\left(\mathbf{F}\cdot\mathbf{P}_{t-1|t-1}\cdot\mathbf{F}'+\mathbf{Q}\right)\mathbf{h}_t\right| \\
&-\frac{1}{2}\sum_{t=1}^{T}\left[\begin{array}{c}
\left(y_t^{+}-\mathbf{a}_t'\cdot\mathbf{x}_t^{*}-\mathbf{h}_t'\cdot\left(\mathbf{F}\hat{\boldsymbol{\xi}}_{t-1|t-1}+\mathbf{C}'\cdot\mathbf{x}_t\right)\right)\cdot \\
\left(\mathbf{h}_t'\left(\mathbf{F}\cdot\mathbf{P}_{t-1|t-1}\cdot\mathbf{F}'+\mathbf{Q}\right)\mathbf{h}_t\right)^{-1}\cdot \\
\left(y_t^{+}-\mathbf{a}_t'\cdot\mathbf{x}_t^{*}-\mathbf{h}_t'\cdot\left(\mathbf{F}\hat{\boldsymbol{\xi}}_{t-1|t-1}+\mathbf{C}'\cdot\mathbf{x}_t\right)\right)'
\end{array}\right]
\end{aligned}
$$

Appendix B: Chow and Lin Regression

In this appendix, we show the Chow and Lin regression model. Chow and Lin assume a true model for the monthly GDP explained by w related series given in matrix notation for the whole sample of T observations: $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u}$ where $\mathbf{V} = E[\mathbf{uu}']$ is the variance-covariance matrix of the error terms. With help of a $\left[\frac{T}{3} \times T\right]$ matrix $\mathbf{C}_D = \frac{1}{3}\left(\begin{pmatrix} 1 & 1 & 1 \end{pmatrix} \quad I_{\frac{T}{3}}\right)$, they transform this true model to match the quarterly observed GDP. The quarterly vector can thus be expressed:

$$\mathbf{y}^{+} = \mathbf{C}_D\mathbf{y} = \mathbf{C}_D\mathbf{X}\boldsymbol{\beta} + \mathbf{C}_D\mathbf{u} = \underset{\left[\frac{T}{3}\times w\right]}{\mathbf{X}^{+}}\boldsymbol{\beta} + \mathbf{u}^{+}$$

where $E\left[\mathbf{u}^{+}\mathbf{u}^{+\prime}\right] = \mathbf{V}^{+} = \mathbf{C}_D\mathbf{V}\mathbf{C}_D'$ and where $\mathbf{X}^{+}$ is a matrix with quarterly average of related series. Chow and Lin looks then for a $\left[T \times \frac{T}{3}\right]$ matrix $\mathbf{A}$ that can fill the gap between quarterly and estimated monthly data such that:

$$\hat{\mathbf{y}} = \mathbf{A}\mathbf{y}^{+}$$

In this search they impose an unbiased estimated monthly series $\hat{\mathbf{y}}$:

$$E\left[\hat{\mathbf{y}} - \mathbf{y}\right] = E\left[\mathbf{A}\left(\mathbf{X}^{+}\boldsymbol{\beta} + \mathbf{u}^{+}\right) - \mathbf{X}\boldsymbol{\beta} - \mathbf{u}\right] = E\left[\left(\mathbf{A}\mathbf{X}^{+} - \mathbf{X}\right)\boldsymbol{\beta}\right] = \mathbf{0}$$

implying that $\mathbf{A}\mathbf{X}^{+} - \mathbf{X} = \mathbf{0}$ and giving then an expression for the difference between the true monthly series and the estimated one $\hat{\mathbf{y}} - \mathbf{y}$:

$$\hat{\mathbf{y}} - \mathbf{y} = \mathbf{A}\mathbf{u}^{+} - \mathbf{u}$$

To find the optimal matrix $\mathbf{A}$, they minimize under the constraint of unbiasedness the trace of the variance-covariance matrix $Cov\left[\hat{\mathbf{y}}\right]$, in fact minimizing the sum of all the variances corresponding to each "observation". The $Cov\left[\hat{\mathbf{y}}\right]$ is:

$$\begin{aligned} E\left[(\hat{\mathbf{y}} - \mathbf{y})^2\right] &= \mathbf{A}E\left[\mathbf{u}^{+}\mathbf{u}^{+\prime}\right]\mathbf{A}' - \mathbf{A}E\left[\mathbf{u}^{+}\mathbf{u}'\right] - E\left[\mathbf{u}\mathbf{u}^{+\prime}\right]\mathbf{A}' + E\left[\mathbf{uu}'\right] \\ &= \mathbf{A}\mathbf{V}^{+}\mathbf{A}' - \mathbf{A}\underbrace{E\left[\mathbf{u}^{+}\mathbf{u}'\right]}_{\mathbf{V}^{+''}} - \underbrace{E\left[\mathbf{u}\mathbf{u}^{+\prime}\right]}_{\mathbf{V}^{''+}}\mathbf{A}' + \mathbf{V} \\ &= \mathbf{A}\mathbf{V}^{+}\mathbf{A}' - \mathbf{A}\mathbf{V}^{+''} - \mathbf{V}^{''+}\mathbf{A}' + \mathbf{V} \end{aligned}$$

They minimize with respect to $\mathbf{A}$ the following Lagrange function with help of a Lagrange multiplier $\mathbf{M}'$:

$$\begin{aligned} L &= \frac{1}{2}tr\left[\mathbf{A}\mathbf{V}^{+}\mathbf{A}' - \mathbf{A}\mathbf{V}^{+''} - \mathbf{V}^{''+}\mathbf{A}' + \mathbf{V}\right] - tr\left[\underset{[w\times T]}{\mathbf{M}}{}'\left(\mathbf{A}\mathbf{X}^{+} - \mathbf{X}\right)\right] \\ L &= \frac{1}{2}tr\left[\mathbf{A}\mathbf{V}^{+}\mathbf{A}'\right] - tr\left[\mathbf{A}\mathbf{V}^{+''}\right] + \frac{1}{2}tr\left[\mathbf{V}\right] - tr\left[\mathbf{X}^{+}\mathbf{M}'\mathbf{A}\right] + tr\left[\mathbf{M}'\mathbf{X}\right] \end{aligned}$$

yielding $\mathbf{A}$:

$$\begin{aligned}\mathbf{A} &= \mathbf{X}\left(\mathbf{X}^{+'}\mathbf{V}^{+^{-1}}\mathbf{X}^{+}\right)^{-1}\mathbf{X}^{+'}\mathbf{V}^{+^{-1}} \\ &\quad +\mathbf{V}^{"+}\mathbf{V}^{+^{-1}}\left[\mathbf{I}_n - \mathbf{X}^{+}\left(\mathbf{X}^{+'}\mathbf{V}^{+^{-1}}\mathbf{X}^{+}\right)^{-1}\mathbf{X}^{+'}\mathbf{V}^{+^{-1}}\right]\end{aligned}$$

The $\hat{\mathbf{y}}$ is then given by the following fitted values: $\hat{\mathbf{y}} = \mathbf{A}\mathbf{y}^{+}$.

$$\begin{aligned}\hat{\mathbf{y}} &= \mathbf{X}\overbrace{\left(\mathbf{X}^{+'}\mathbf{V}^{+^{-1}}\mathbf{X}^{+}\right)^{-1}\mathbf{X}^{+'}\mathbf{V}^{+^{-1}}\mathbf{y}^{+}}^{\hat{\boldsymbol{\beta}}_{GLS}} \\ &\quad +\mathbf{V}^{"+}\mathbf{V}^{+^{-1}}\underbrace{\left[\mathbf{I}_{\frac{T}{3}} - \mathbf{X}^{+}\left(\mathbf{X}^{+'}\mathbf{V}^{+^{-1}}\mathbf{X}^{+}\right)^{-1}\mathbf{X}^{+'}\mathbf{V}^{+^{-1}}\right]\mathbf{y}^{+}}_{\hat{\mathbf{u}}^{+}} \\ \hat{\mathbf{y}} &= \mathbf{X}\hat{\boldsymbol{\beta}}_{GLS}+\left(\mathbf{V}^{"+}\mathbf{V}^{+^{-1}}\right)\hat{\mathbf{u}}^{+} = \mathbf{X}\hat{\boldsymbol{\beta}}_{GLS}+\boldsymbol{\Lambda}\hat{\mathbf{u}}^{+}\end{aligned}$$

The monthly series is computed by the third of all the elements of the vector $\hat{\mathbf{y}}$.

Appendix C: Comparison of Log-Likelihood Functions

In this appendix, we show that the Kalman filter and the Chow and Lin regression yield the same estimates by maximum likelihood. Assume this structural equation:

$$y_t = \mathbf{x}_t'\boldsymbol{\beta} + u_t, \text{ for } t = 1, \ldots, T$$

State vector:

$$\boldsymbol{\xi}_t = \begin{pmatrix} y_t - \mathbf{x}_t'\boldsymbol{\beta} \\ y_{t-1} - \mathbf{x}_{t-1}'\boldsymbol{\beta} \\ y_{t-2} - \mathbf{x}_{t-2}'\boldsymbol{\beta} \end{pmatrix}, E(u_t^2) = \sigma^2$$

State equation:

$$\boldsymbol{\xi}_t = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \boldsymbol{\xi}_{t-1} + \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} u_t \\ u_{t-1} \\ u_{t-2} \end{pmatrix}$$

Measurement equation:

$$y_t^+ = \mathbf{a}_t'\mathbf{x}_t^* + \mathbf{h}_t'\boldsymbol{\xi}_t \text{ or } y_t^+ = \mathbf{a}_t'\left(\mathbf{x}_t + \mathbf{x}_{t-1} + \mathbf{x}_{t-2}\right) + \mathbf{h}_t'\boldsymbol{\xi}_t$$

where:

$$\begin{aligned}
\mathbf{x}_t^* &= \sum_{j=t-2}^{t} x_j \\
\mathbf{h}_t' &= \begin{pmatrix} 0 & 0 & 0 \end{pmatrix} \text{ and } \mathbf{a}_t' = 0 \text{ for } t = 1, 2, 4, 5, 7, \ldots, T-1 \\
\mathbf{h}_t' &= \begin{pmatrix} 1 & 1 & 1 \end{pmatrix} \text{ and } \mathbf{a}_t' = \boldsymbol{\beta}' \text{ for } t = 3, 6, 9, \ldots, T
\end{aligned}$$

Assume:

$$\hat{\boldsymbol{\xi}}_{t|t-1} = \begin{pmatrix} \hat{y}_{t|t-1} - \mathbf{x}_t'\hat{\boldsymbol{\beta}} \\ \hat{y}_{t-1|t-1} - \mathbf{x}_{t-1}'\hat{\boldsymbol{\beta}} \\ \hat{y}_{t-2|t-1} - \mathbf{x}_{t-2}'\hat{\boldsymbol{\beta}} \end{pmatrix}$$

$$\mathbf{P}_{t|t-1} = \begin{pmatrix} \sigma^2 & 0 & 0 \\ 0 & \sigma^2 & 0 \\ 0 & 0 & \sigma^2 \end{pmatrix}$$

Maximum likelihood function for this Kalman filter:

$$\begin{aligned}
\sum_{t=1}^{\frac{T}{3}} \ln f\left(y_t^+\right) &= -\frac{T}{6}\ln(2\pi) - \frac{T}{12}\ln\left(3\sigma^2\right) \\
&\quad -\frac{1}{2}\sum_{\tau=1}^{\frac{T}{3}} \left[\begin{array}{c} \left(y_{3\tau}^+ - \mathbf{a}_{3\tau}'\mathbf{x}_{3\tau}^* - \mathbf{h}_{3\tau}' \cdot \hat{\boldsymbol{\xi}}_{3\tau|3\tau-1}\right) \\ \left(3\sigma^2\right)^{-1}\left(y_{3\tau}^+ - \mathbf{a}_{3\tau}'\mathbf{x}_{3\tau}^* - \mathbf{h}_{3\tau}' \cdot \hat{\boldsymbol{\xi}}_{3\tau|3\tau-1}\right)' \end{array} \right]
\end{aligned}$$

$$
\begin{aligned}
\sum_{t=1}^{\frac{T}{3}} \ln f\left(y_t^{+}\right) &= -\frac{T}{6}\ln(2\pi) - \frac{T}{12}\ln\left(3\sigma^2\right) \\
&\quad -\frac{1}{6\sigma^2}\sum_{\tau=1}^{\frac{T}{3}} \left(\begin{array}{c} y_{3\tau}^{+} \\ -\boldsymbol{\beta}'\left(\mathbf{x}_{3\tau}+\mathbf{x}_{3\tau-1}+\mathbf{x}_{3\tau-2}\right) \\ -\left(\begin{array}{c} \hat{y}_{3\tau|3\tau-1} - \mathbf{x}_{3_\tau}\hat{\boldsymbol{\beta}} \\ +\hat{y}_{3\tau-1|3\tau-1} - \mathbf{x}_{3\tau-1}\hat{\boldsymbol{\beta}} \\ +\hat{y}_{3\tau-2|3\tau-1} - \mathbf{x}_{3\tau-2}\hat{\boldsymbol{\beta}} \end{array} \right) \end{array} \right)^2
\end{aligned}
$$

$$
\begin{aligned}
\sum_{t=1}^{\frac{T}{3}} \ln f\left(y_t^{+}\right) &= -\frac{T}{6}\ln(2\pi) - \frac{T}{12}\ln\left(3\sigma^2\right) \\
&\quad -\frac{1}{6\sigma^2}\sum_{\tau=1}^{\frac{T}{3}} \left(\begin{array}{c} y_{3\tau}^{+} \\ -\left(\begin{array}{c} \hat{y}_{3\tau|3\tau-1} \\ +\hat{y}_{3\tau-1|3\tau-1} \\ +\hat{y}_{3\tau-2|3\tau-1} \end{array} \right) \end{array} \right)^2
\end{aligned}
$$

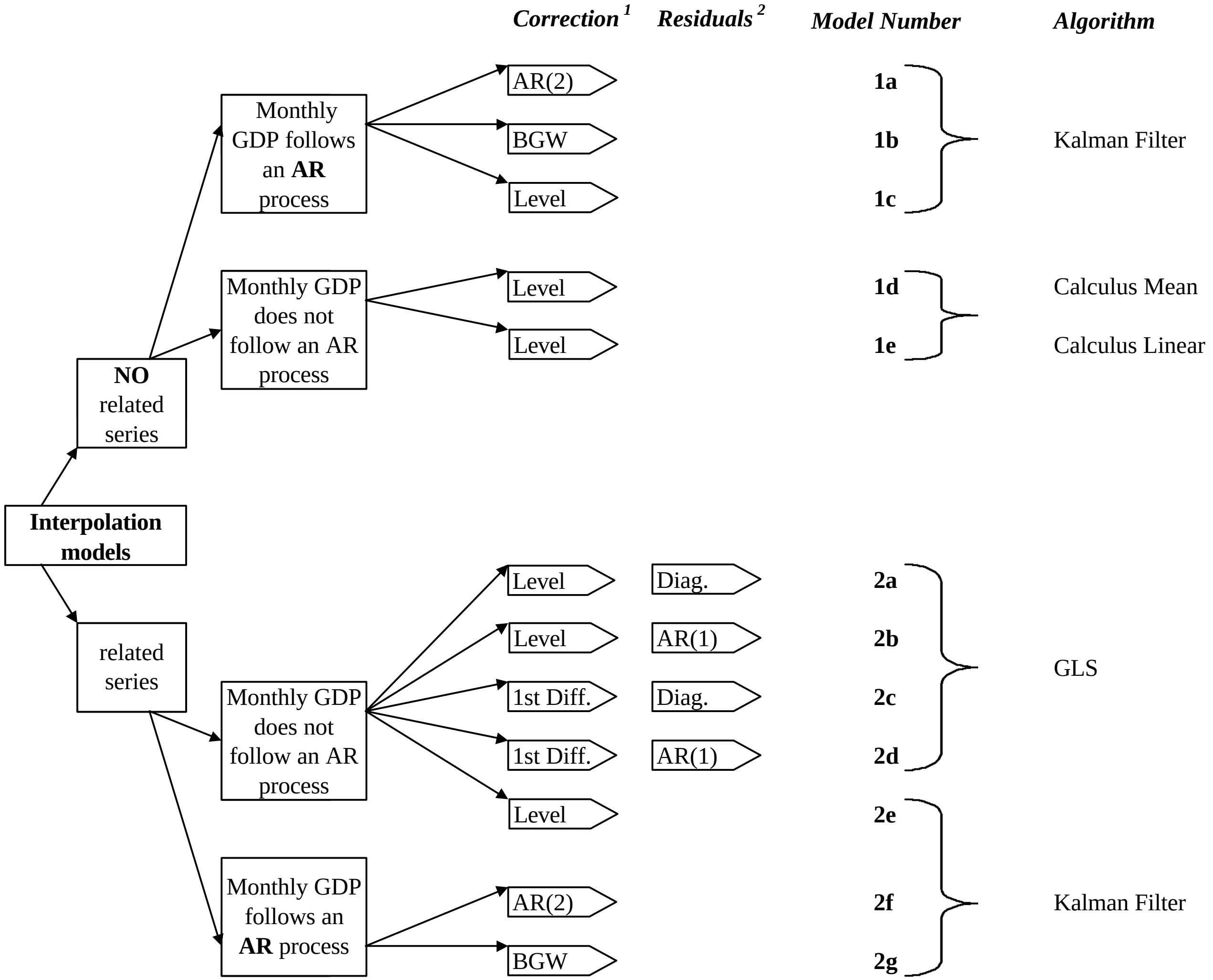
and this is the log-likelihood function for the Chow and Lin regression where for $\mathbf{y}^{+}$:

$$
y_t^{+} \sim N\left(\underset{E[u^{+}]}{\hat{\boldsymbol{\beta}}'\mathbf{x}_t^{*}}, \underset{V[u^{+}]}{3\sigma^2}\right)
$$

or

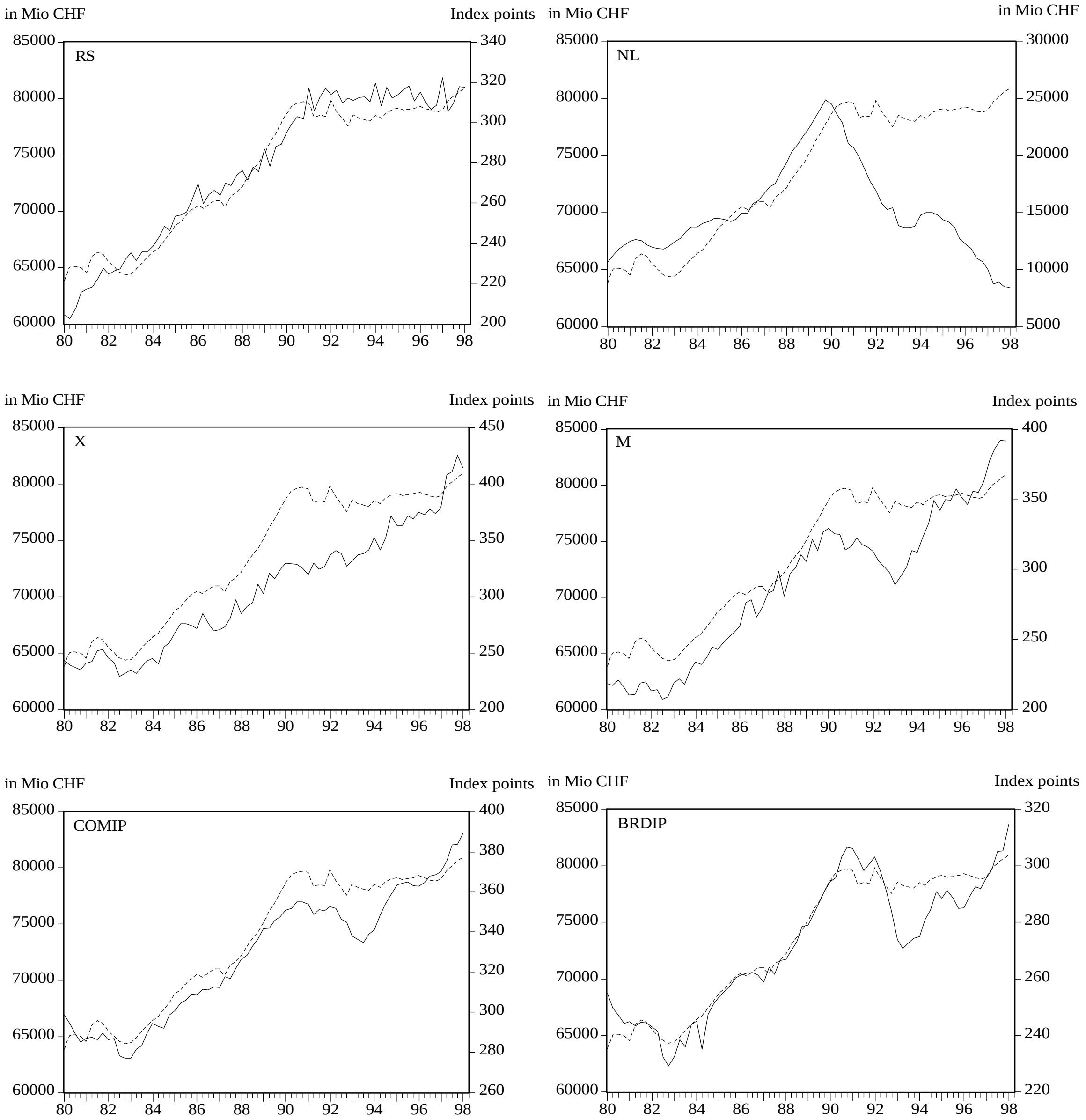
$$
y_t^{+} \sim N \quad \underbrace{(\hat{y}_{t|t-1} + \hat{y}_{t-1|t-1} + \hat{y}_{t-2|t-1}}_{\substack{\hat{\boldsymbol{\beta}}'\mathbf{x}_t + \hat{\boldsymbol{\beta}}'\mathbf{x}_{t-1} + \hat{\boldsymbol{\beta}}'\mathbf{x}_{t-2} = \\ \hat{\boldsymbol{\beta}}'\left(\mathbf{x}_t + \mathbf{x}_{t-1} + \mathbf{x}_{t-2}\right) = \hat{\boldsymbol{\beta}}'\mathbf{x}_t^{*}}} \quad , 3\sigma^2)
$$

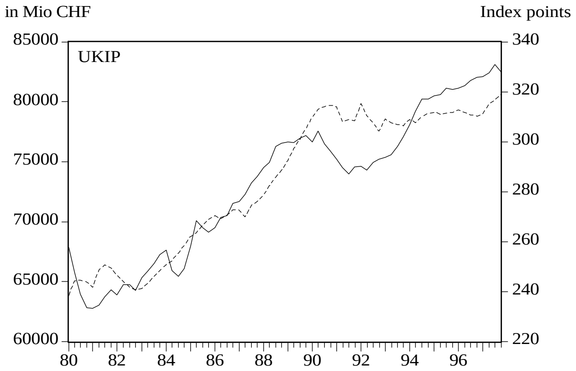
Figure 1 - Overview of Interpolation Models



[1]AR(2) stands for an AR(1) process in first difference rewritten as an AR(2) in levels; BGW means correction according to Bernanke, Gertler and Watson [1997]; 1st Diff. uses a first difference operator. [2]Diag. indicates no autocorrelation; AR(1) stands for residuals following an AR(1) process.

Figure 2 - GDP (-) and Related Series (--)[1,2]





[1] GDP = Quarterly Gross Domestic Product; RS = Value of retail sales; NL = Level of not utilized construction loans; X = Exports volume; M = Imports volume; BRDIP = Industrial production in Germany; UKIP = Industrial production in UK; COMIP = Composite index of industrial productions.

[2] Source: Datastream.

Figure 3 - Quarterly GDP and Interpolated Series with Pattern

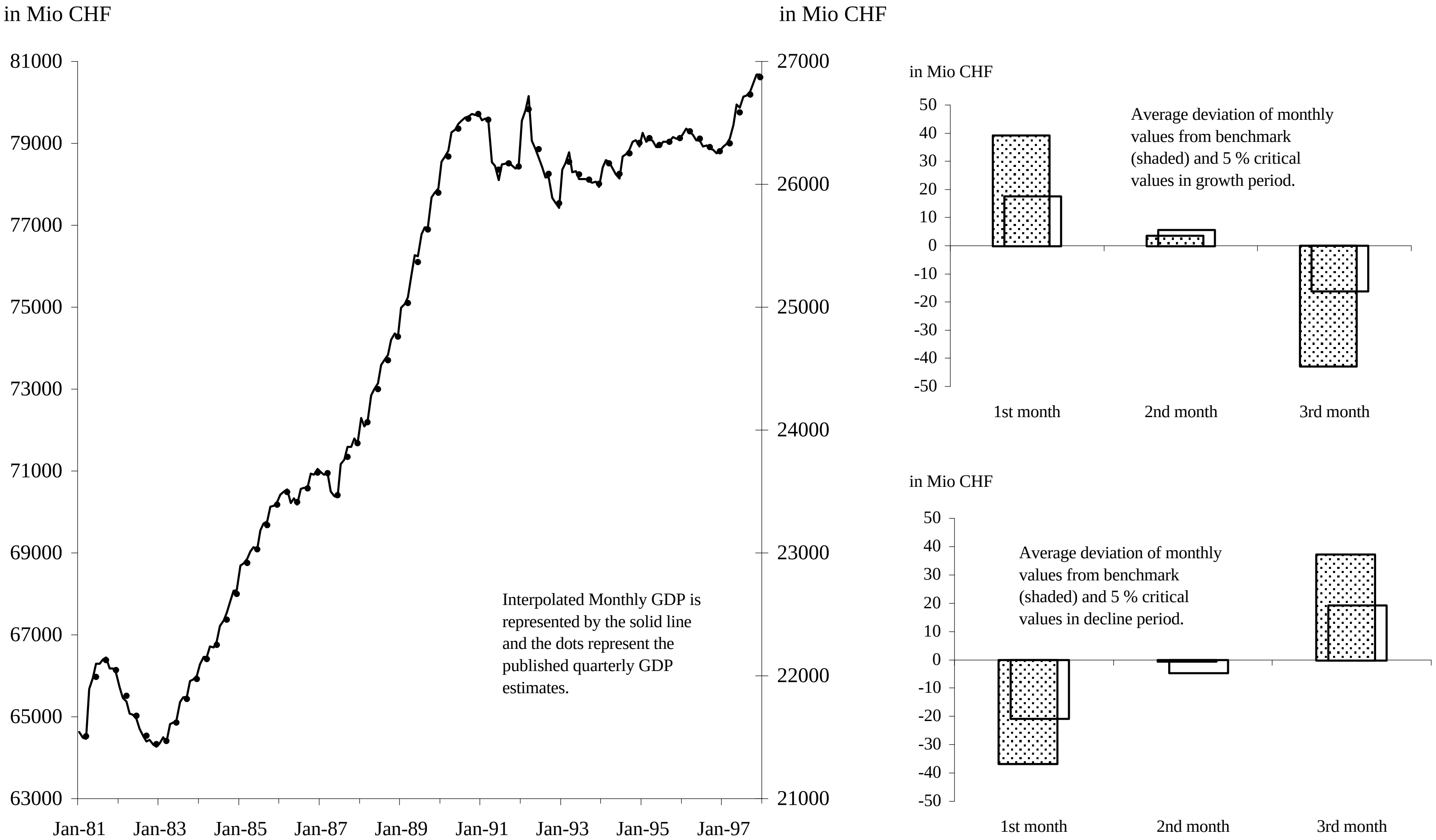


Figure 4 - Monthly GDP 81-97

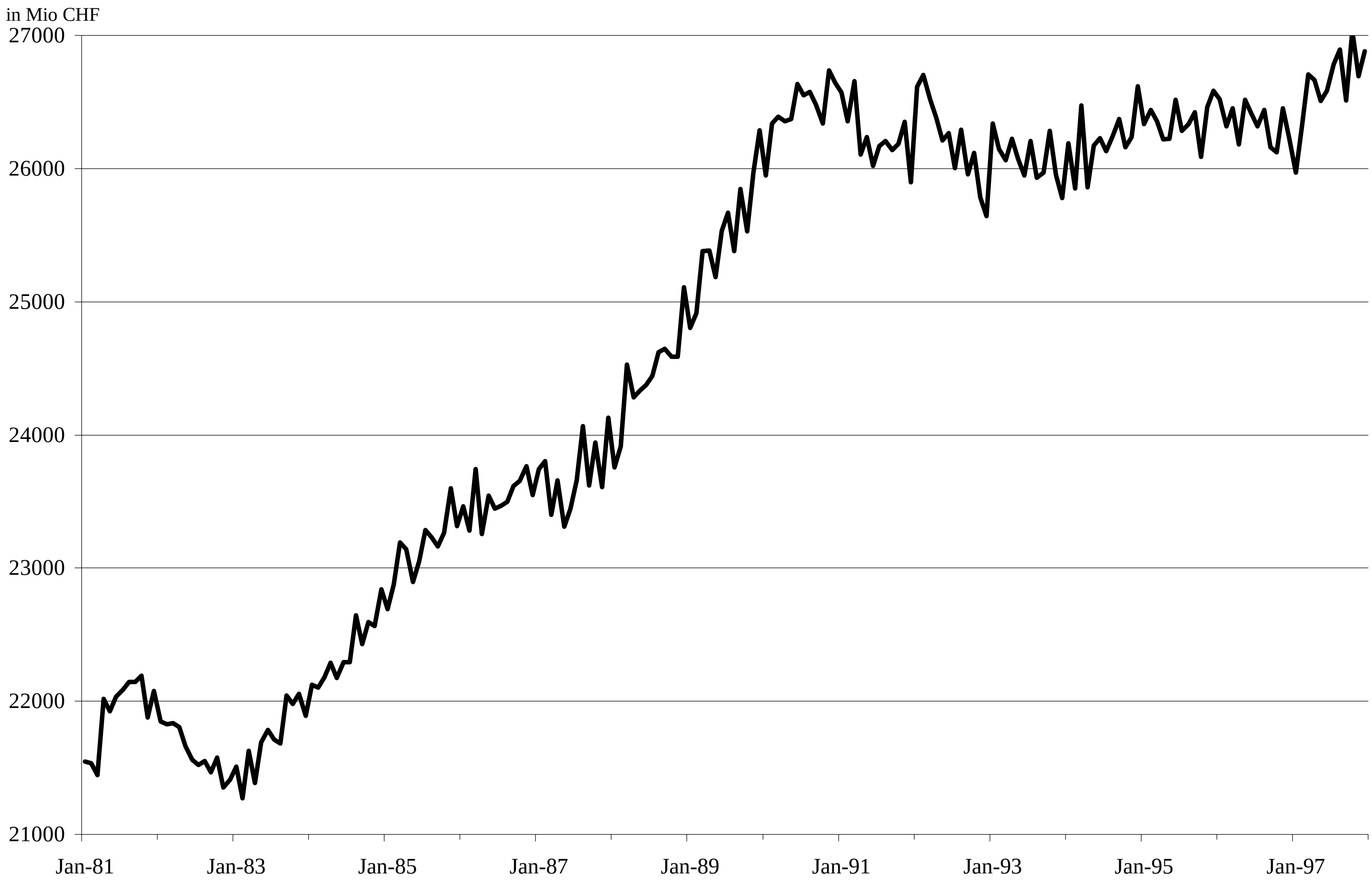


Table 1 - Descriptive Statistics of Observed Time Series[1]

	GDP[2,3]	RS	NL	X	M	BRDIP	UKIP	COMIP
mean	1.330	3.198	-0.966	4.087	4.289	1.402	1.310	1.516
st.dev.	2.947	46.289	21.386	49.032	51.113	21.732	13.170	12.163
AR(1)[4]	0.252	-0.653 *	0.249 *	-0.569 *	-0.609 *	-0.428 *	-0.220 *	-0.323 *
JB test	0.050	77.884 *	199.177 *	53.174 *	85.518 *	1370.582 *	6.710 *	68.533 *
ADF	-4.379 *	-11.043 *	-2.981 *	-7.913 *	-8.064 *	-6.261 *	-5.298 *	-5.632 *
GDP,SERIES(-4)[5]		0.086	0.361	-0.082	0.060	0.098	-0.117	0.069
GDP,SERIES(-3)		0.001	0.296	0.103	0.156	0.031	0.030	0.049
GDP,SERIES(-2)		0.124	0.372	0.147	0.175	0.118	0.092	0.176
GDP,SERIES(-1)		-0.006	0.304	0.183	0.202	0.336	0.168	0.409
GDP,SERIES(0)		0.089	0.232	0.261	0.088	0.248	0.050	0.267
GDP,SERIES(1)		0.126	0.235	0.254	0.257	0.352	-0.012	0.302
GDP,SERIES(2)		0.123	0.145	-0.019	-0.053	0.212	-0.088	0.187
GDP,SERIES(3)		0.022	0.164	-0.073	0.035	0.143	-0.172	0.089
GDP,SERIES(4)		0.191	0.027	-0.057	-0.186	-0.076	-0.064	-0.120

Note:

[1]Annualized statistical figures are calculated for quarterly growth rates of GDP and for monthly growth rates for all other variables.

[2]GDP = Gross domestic product; RS = Value of retail sales; NL = Level of not utilized construction loans; X = Exports volume; M = Imports volume; BRDIP = Industrial production in Germany; UKIP = Industrial productin in UK; COMIP = Composite index of IP.

[3]All variables except COMIP are seasonally adjusted.

[4] ** = significant at 5 % level; * = significant at 1 % level, for the t-test of a zero coefficient on an AR(1) process, for the Jarque-Bera (JB) test of normal distribution, and for the augmented Dickey-Fuller test of unit root (ADF).

[5]Cross-correlation of leads and lags of quarterly growth rates of related series with quarterly GDP growth rate.

Table 2 - Cointegration Test of GDP and RS[1]

H_0	H_a	Eigenvalue[2]	LR[3]	1%	5%
0	**2**	0.2840	25.1144 *	20.040	15.4100
1	**2**	0.0244	1.7254	6.65	3.7600
0	**1**	0.2840	23.3891 *	6.65	3.7600

Note:

[1]Cointegration tests are performed with quarterly level data. GDP = Gross domestic product; RS = Value of retail sales.

[2]Tests are run assuming linear trend in data and an intercept in the cointegrating equation and in the vector autoregression. Two lags are included.

[3] ** = significant at 5 % level; * = significant at 1 % level.

Table 3 - Interpolation Results[1,2]

Model	1	2	3	4	5
Class, Related Series	1a	1c	2a,COMIP	2a;NL	2a;X,M,RS,NL
AIC	8.0484	10.0798	14.6576	17.3080	13.0627
log L.	-562.5559	-573.4082	-636.5871	-733.3278	-575.3729
mean	1.3179	1.3276	1.1978	1.2523	1.2675
std	4.3039	5.1094	9.9281	5.2206	14.1292
AR(1)[3]	0.2154 *	0.0595	-0.3705 *	-0.0167	-0.5323 *
JB test	308.0648 *	15.5904 *	12.0892 **	192.1336 *	2.7653
ADF test	-5.4689 *	-5.9574 *	-5.7885 *	-5.3002 *	-6.4013 *
MSE w. benchmark[4]	3307.38	4146.21	18659.04	7123.27	32328.09
MSE A->Q	196862.53	144095.19	206864.82	208184.95	248244.08

Model	6	7	8	9	10
Class, Related Series	2b;COMIP	2b;NL	2b;X,M,RS,NL	2c;COMIP	2c;NL
AIC	15.8917	17.3255	13.3458	15.6284	14.4199
log L.	-681.6513	-733.9669	-585.7091	-677.9689	-627.1078
mean	1.2868	1.3038	1.4108	1.2961	1.3025
std	6.4232	3.4897	12.2157	3.6553	3.5096
AR(1)[3]	-0.1438 **	0.7343 *	-0.5350 *	0.6266 *	0.7217 *
JB test	17.9339 *	16.8364 *	0.5234	6.5835 **	15.9277 *
ADF test	-5.5394 *	-4.2072 *	-6.8464 *	-4.3023 *	-4.1666 *
MSE w. benchmark[4]	8884.46	4245.59	23601.90	4394.56	4287.16
MSE A->Q	129891.30	97166.50	221468.03	139871.77	102001.75

Model	11	12	13	14	15
Class, Related Series	2c;X,M,RS,NL	2d;COMIP	2d;NL	2d;X,M,RS,NL	2e;COMIP
AIC	14.3756	15.2662	14.3024	14.0856	11.4703
log L.	-623.5207	-664.8914	-622.8601	-614.8347	-569.7796
mean	1.3375	1.2685	1.2758	1.3468	1.3820
std	4.6631	6.8035	4.6896	7.9571	9.9682
AR(1)[3]	0.1387 **	-0.2316 *	0.0975	-0.3632 *	-0.3851 *
JB test	9.6226 *	13.4503 *	332.6322 *	15.0818 *	17.7948 *
ADF test	-5.5733 *	-4.9152 *	-4.9212 *	-5.6324 *	-5.6695 *
MSE w. benchmark[4]	5674.41	10947.86	6696.91	11778.87	19095.04
MSE A->Q	79448.95	120193.62	151188.26	141821.70	259232.93

Model	16	17	18	19	20
Class, Related Series	2e;NL	2e;X,M,RS,NL	2f;COMIP	2f;NL	2f;X,M,RS,NL
AIC	13.9672	10.9534	8.0537	8.0498	8.0407
log L.	-707.0603	-564.4899	-562.5552	-562.5514	-562.5456
mean	1.2766	1.3828	1.3181	1.2846	1.2839
std	5.2332	12.1565	4.3022	4.3011	4.5758
AR(1)[3]	-0.0188	-0.5220 *	0.2155 *	0.2235 *	0.1080
JB test	216.2845 *	2.5033	289.2578 *	306.5034 *	156.8350 *
ADF test	-5.2316 *	-6.2078 *	-5.4785 *	-5.4467 *	-5.5563 *
MSE w. benchmark[4]	7152.71	23682.66	3304.75	3282.52	3701.94
MSE A->Q	210478.57	230738.92	488844.80	147968.33	163693.19

Note: [1]GDP = Gross domestic product; RS = Value of retail sales; NL = Level of not utilized construction loans;

X = Exports volume; M = Imports volume; BRDIP = Industrial production in Germany; UKIP = Industrial production in UK;

COMIP = Composite index of IP. All estimations include a constant, models 2c and 2d transform time trend to constant.

[2]The descriptive statistics are for growth rates of the interpolated GDP for 81-97.

[3] ** = significant at 5 % level; * = significant at 1 % level, for all the tests.

[4]Level MSE with benchmark 1e is for the period 81-97 and level MSE A->Q is for the whole period 82-96.

Table 4 - Interpolated GDP[1,2]

M1 81	21547.19	M1 84	22122.76	M1 87	23741.70	M1 90	25950.00	M1 93	26339.71	M1 96	26518.88
M2 81	21533.94	M2 84	22101.38	M2 87	23803.74	M2 90	26338.53	M2 93	26149.14	M2 96	26319.14
M3 81	21445.75	M3 84	22180.70	M3 87	23399.89	M3 90	26390.93	M3 93	26062.83	M3 96	26454.22
	64526.88		*66404.84*		*70945.33*		*78679.46*		*78551.69*		*79292.24*
M4 81	22019.38	M4 84	22288.97	M4 87	23659.96	M4 90	26353.83	M4 93	26225.01	M4 96	26180.06
M5 81	21924.48	M5 84	22173.38	M5 87	23309.09	M5 90	26370.57	M5 93	26071.01	M5 96	26518.21
M6 81	22034.11	M6 84	22294.04	M6 87	23444.59	M6 90	26636.30	M6 93	25949.31	M6 96	26416.75
	65977.97		*66756.39*		*70413.63*		*79360.69*		*78245.33*		*79115.02*
M7 81	22086.77	M7 84	22293.96	M7 87	23656.61	M7 90	26551.29	M7 93	26208.54	M7 96	26317.14
M8 81	22146.44	M8 84	22646.49	M8 87	24066.38	M8 90	26577.40	M8 93	25932.67	M8 96	26438.84
M9 81	22146.51	M9 84	22427.77	M9 87	23619.80	M9 90	26478.80	M9 93	25968.73	M9 96	26159.72
	66379.72		*67368.22*		*71342.79*		*79607.48*		*78109.94*		*78915.70*
M10 81	22190.22	M10 84	22592.54	M10 87	23942.52	M10 90	26339.18	M10 93	26281.98	M10 96	26123.76
M11 81	21879.02	M11 84	22566.74	M11 87	23608.97	M11 90	26735.52	M11 93	25952.01	M11 96	26452.74
M12 81	22076.78	M12 84	22839.18	M12 87	24130.27	M12 90	26642.29	M12 93	25779.11	M12 96	26234.10
	66146.02		*67998.45*		*71681.76*		*79716.98*		*78013.10*		*78810.60*
M1 82	21847.19	M1 85	22692.26	M1 88	23754.84	M1 91	26570.56	M1 94	26189.55	M1 97	25970.37
M2 82	21828.27	M2 85	22877.37	M2 88	23911.03	M2 91	26354.49	M2 94	25849.91	M2 97	26326.17
M3 82	21836.34	M3 85	23192.12	M3 88	24526.80	M3 91	26655.18	M3 94	26473.88	M3 97	26708.77
	65511.80		*68761.75*		*72192.67*		*79580.23*		*78513.34*		*79005.32*
M4 82	21805.96	M4 85	23141.78	M4 88	24283.51	M4 91	26104.60	M4 94	25860.55	M4 97	26663.08
M5 82	21659.71	M5 85	22895.66	M5 88	24334.66	M5 91	26236.97	M5 94	26172.53	M5 97	26507.56
M6 82	21559.15	M6 85	23053.29	M6 88	24376.48	M6 91	26021.44	M6 94	26226.25	M6 97	26585.48
	65024.82		*69090.73*		*72994.65*		*78363.01*		*78259.34*		*79756.12*
M7 82	21523.04	M7 85	23286.67	M7 88	24443.01	M7 91	26166.88	M7 94	26131.31	M7 97	26782.24
M8 82	21550.88	M8 85	23230.96	M8 88	24622.28	M8 91	26205.93	M8 94	26249.00	M8 97	26895.27
M9 82	21466.62	M9 85	23163.66	M9 88	24644.74	M9 91	26140.26	M9 94	26370.52	M9 97	26514.14
	64540.54		*69681.29*		*73710.03*		*78513.07*		*78750.83*		*80191.65*
M10 82	21575.79	M10 85	23263.82	M10 88	24587.87	M10 91	26184.97	M10 94	26160.96	M10 97	27035.96
M11 82	21351.93	M11 85	23598.24	M11 88	24587.23	M11 91	26350.56	M11 94	26235.31	M11 97	26692.67
M12 82	21411.29	M12 85	23314.73	M12 88	25109.68	M12 91	25896.47	M12 94	26618.72	M12 97	26882.25
	64339.01		*70176.79*		*74284.78*		*78432.01*		*79015.00*		*80610.89*
M1 83	21510.41	M1 86	23465.24	M1 89	24803.75	M1 92	26614.96	M1 95	26334.31		
M2 83	21272.26	M2 86	23281.20	M2 89	24914.50	M2 92	26701.74	M2 95	26438.23		
M3 83	21626.39	M3 86	23743.55	M3 89	25381.97	M3 92	26519.75	M3 95	26354.10		
	64409.06		*70489.99*		*75100.23*		*79836.45*		*79126.64*		
M4 83	21385.29	M4 86	23254.29	M4 89	25386.43	M4 92	26382.97	M4 95	26220.37		
M5 83	21689.72	M5 86	23544.85	M5 89	25186.00	M5 92	26210.13	M5 95	26224.09		
M6 83	21784.81	M6 86	23446.60	M6 89	25533.21	M6 92	26267.09	M6 95	26516.54		
	64859.83		*70245.74*		*76105.64*		*78860.19*		*78961.00*		
M7 83	21713.98	M7 86	23467.64	M7 89	25667.20	M7 92	26001.89	M7 95	26283.06		
M8 83	21681.48	M8 86	23499.52	M8 89	25378.94	M8 92	26293.35	M8 95	26333.46		
M9 83	22044.53	M9 86	23614.40	M9 89	25847.39	M9 92	25957.61	M9 95	26425.39		
	65439.99		*70581.57*		*76893.53*		*78252.86*		*79041.90*		
M10 83	21981.35	M10 86	23652.42	M10 89	25529.57	M10 92	26116.96	M10 95	26088.01		
M11 83	22055.13	M11 86	23764.32	M11 89	25970.32	M11 92	25782.27	M11 95	26460.65		
M12 83	21890.95	M12 86	23549.15	M12 89	26289.12	M12 92	25643.24	M12 95	26582.78		
	65927.43		*70965.89*		*77789.01*		*77542.47*		*79131.45*		

[1] Quarterly values are in italic.

[2] in Mio CHF.