University of Groningen, Research School SOM
Research Report **05F10**

Time Series Factor Analysis with an Application to Measuring Money*

Paul D. Gilbert
Department of Monetary and Financial Analysis, Bank of Canada,
234 Wellington Street, Ottawa, Canada, K1A 0G9
pgilbert@bank-banque-canada.ca

Erik Meijer
Department of Econometrics, University of Groningen,
PO Box 800, 9700 AV Groningen, The Netherlands
e.meijer@eco.rug.nl

November 2005

SOM-theme F Interactions between consumers and firms

Electronic version: http://som.rug.nl

* This paper was prepared in part while the second author was at the Bank of Canada. The views expressed in this paper are those of the authors. No responsibility for them should be attributed to the Bank of Canada. We would like to thank Tom Wansbeek for comments on an earlier version of this paper.

Abstract

Time series factor analysis (TSFA) and its associated statistical theory is developed. Unlike dynamic factor analysis (DFA), TSFA obviates the need for explicitly modeling the process dynamics of the underlying phenomena. It also differs from standard factor analysis (FA) in important respects: the factor model has a nontrivial mean structure, the observations are allowed to be dependent over time, and the data does not need to be covariance stationary as long as differenced data satisfies a weak boundedness condition. The effects on the estimation of parameters and prediction of the factors is discussed.

The statistical properties of the factor score predictor are studied in a simulation study, both over repeated samples and within a given sample. Some apparent anomalies are found in simulation experiments and explained analytically.

The main empirical result from this simulation is that, contrary to what is usually expected in cross-sectional factor analysis, the sampling variability in the (time-invariant) parameters, which is $O_p(T^{-1/2})$, accounts for most of the prediction errors, and the fundamental inability to estimate the factors consistently, which accounts for an $O_p(1)$ term in the prediction error, turns out to have only a very small impact for this type of data.

The application motivating this research is the desire to find good measures of important underlying macro-economic phenomena affecting the financial side of the economy. Technological innovations in the financial industry pose major problems for the measurement of monetary aggregates. The TSFA estimation methodology proposed in this paper provides a way to obtain new measures that are more robust to the effects of financial innovations. The example application uses the general ideas laid out in Gilbert and Pichette (2003), but explores the improved estimation methods of TSFA. What was considered an important difficulty in that paper is now understood and shown not to be a serious problem. The approach has considerable promise for replacing the monetary aggregates.

1 Introduction

Standard factor analysis (FA) does not work directly with typical macro-economic time series because the characteristics of the data usually conflict with the assumptions. FA (see for example Wansbeek and Meijer, 2000, chap. 7), was developed for cross-sectional data where the assumptions are often reasonable. Most notably, FA theory assumes observations are independent and identically distributed (i.i.d.). Macro-economic data typically trend upwards and are serially dependent, so the i.i.d. assumption is violated.

Furthermore, most FA applications assume that intercepts are uninteresting free parameters, which implies that sample means can be subtracted and the centered data treated as mean zero and i.i.d. Time series applications have two important reasons why means are of interest and intercepts should be restricted to zero. First, the intuition of the interpretation is clearer: if the factors are zero, the explained phenomena are also zero. Second, macro-economic variables are often interpreted in growth rates, and the mean affects the magnitude of growth rates.

Dynamic factor analysis (DFA), often based on state-space models, was developed to address these differences (see, e.g., Watson and Engle, 1983; Harvey, 1989; Hamilton, 1994). State space models specify how the observed variables are related to the factors (the states) and also specify a dynamic model for the factors. Molenaar (1985) proposed a DFA model with a seemingly different model structure, but his model can be rewritten in an equivalent state-space form.

The drawback of modeling factor dynamics is that a substantive model of the factors must be specified. Consequently, parameter estimates of the measurement process, and resulting factor "predictions" depend critically on the specified dynamic factor model. This is often undesirable because differences between economic models may be exaggerated or blurred by the resulting data measurement differences. The possibility of estimating parameters and predicting factor scores under minimal assumptions about factor dynamics is explored below. The name *time series factor analysis* (TSFA) is used to distinguish these techniques from DFA. This paper develops the TSFA estimation methodology for integrating time series data. Corrections are also included to accommodate nonzero means.

An important field of research on factor models for time series is the class of (static and dynamic) *approximate factor models*. These models are suitable for time series data with many series, with relatively few underlying factors. The typical application is asset return data, with both a large number of time points and a large number of assets

considered. The large number of series means that asymptotics can be used in which both dimensions of the data matrix are diverging to infinity. Model assumptions such as the uncorrelatedness of error terms can then be relaxed and both the parameters and the underlying factors can be consistently estimated. See, e.g., Chamberlain and Rothchild (1983), Forni et al. (2000), Bai and Ng (2002), Bai (2003), and Stock and Watson (2005) for these models. In contrast, TSFA is suitable for a fixed (relatively small) number of series and therefore relies on somewhat stronger model assumptions.

TSFA should be useful when the researcher does both measurement and modeling, because specific assumptions about factor dynamics are usually much more fragile than the assumption that factors exist. With TSFA the factors can be measured before modeling their dynamics. However, TSFA may be especially important where one group (e.g., a statistics agency or central bank) measures data for many researchers to use.

Geweke (1977) also defined a factor analysis model for a multivariate time series without explicitly specifying the dynamic model for the factors, but he assumed covariance stationarity. This allowed estimation of parameters in the frequency domain. In contrast, TSFA does not assume covariance stationarity and estimation is in the time domain.

TSFA is also closely related to the "*P*-technique", proposed by Cattell (1943) and Cattell et al. (1947), which applied standard FA to multivariate time series. In the development of *P*-technique no explicit assumptions were stated and practices were used for which the methodological basis is questionable. First, the data were not de-trended. Estimators are shown below to have desirable statistical properties such as consistency after de-trending, which may not be the case otherwise. Second, a substantive model was estimated in an attempt to accommodate the dynamic process. This was done by including exogenous variables and deterministic functions of time that were treated as additional indicators, and by using a matrix of the largest cross-correlations rather than an ordinary correlation matrix. That is, if x and y are two observed variables, $\mathrm{Corr}(x_t, y_t)$ was replaced by $\mathrm{Corr}(x_t, y_s)$, where s is such that the absolute value of this correlation is maximized. The *P*-technique, and especially this implementation, has been heavily criticized by Anderson (1963) and Holtzman (1962). TSFA does not include exogenous variables and deterministic functions of time, and only uses a proper covariance matrix (or correlation matrix). Furthermore, data is de-trended by differencing and weak assumptions under which TSFA gives consistent estimates are explicitly stated below.

Finally, this paper is related to Spanos (1984), both in terms of methodology and application. He first estimated a FA model from first differences of a multivariate time series, and then predicted the factor scores, which he used in a subsequent

analysis of an economic model. Explicit assumptions are missing, but i.i.d. appears to be assumed. After model specification, he re-estimated the complete model in a state-space form without a dynamic factor relationship. In the application to measuring money, he used only one factor and presumed that this would represent liquidity, as he thought that this was the most common aspect of the various indicators. In contrast, below weak assumptions are stated explicitly and subsequent economic models are not discussed, the properties of the estimators and factor score predictors are studied through simulation, and in the application a number of different choices (number of factors, construction of the indicators) are made.

In the example application in section 5, TSFA is illustrated as a way to link measured data, also called *indicators* (currency and deposit balances), to the *factors* which are the underlying phenomena of interest (the intended use of money for transactions and savings). Historically, monetary aggregates have been used to measure activity in the financial side of the economy. Their ability to predict economic activity and inflation has been subject to much debate. The problems with these traditional measures are discussed in Gilbert and Pichette (2002, 2003), and in many of the references cited in those papers. While these traditional measures are now largely unused, we hope that a better understanding of the financial side of the economy would be useful, and ultimately lead to models which are better for policy and prediction. Better measurement is a necessary first step in this process.

The organization of the paper is as follows. Section 2 defines the TSFA model, states weak regularity assumptions that will be used, and discusses statistical properties of estimators. Section 3 develops theory for factor score prediction in the given context. Section 4 gives a Monte Carlo analysis of the techniques, with a sample size and data as might be expected in many macro-economics problems. Section 5 gives an example using the application motivating this research, extracting factors from Canadian money data. Sections 6 discusses the sensitivity of the results to the selected sample. Finally, section 7 summarizes the results and discusses outstanding issues.

2 Time series factor analysis (TSFA)

The k unobserved processes of interest (the *factors*) for a sample of T time periods will be indicated by ξ_{it}, $t = 1, \ldots, T$, $i = 1, \ldots, k$. The M observed processes (the *indicators*) will be denoted by y_{it}, $t = 1, \ldots, T$, $i = 1, \ldots, M$. The factors and indicators for period t are collected in the (column) vectors ξ_t and y_t, respectively. It is assumed there is a measurement model relating the indicators to the factors given by

$$y_t = \alpha + B\xi_t + \varepsilon_t, \tag{1}$$

where α is an M-vector of intercept parameters, B is an $M \times k$ matrix parameter of *factor loadings* or simply *loadings*, and ε_t is a random M-vector of measurement errors, disturbances, and unique or idiosyncratic factors. In the example application it is assumed that $\alpha = 0$ but the theory is developed for the general case.

Equation (1) is a standard FA model except that indicators are indexed by time and intercepts are explicitly included, whereas in FA means are usually subtracted. The fact that data are time series is important mainly because economic data are typically growing and thus not covariance stationary. Other than this, the sequential order of the data is irrelevant in TSFA as opposed to DFA.

FA is usually applied to cross-sectional data where it is reasonable to assume i.i.d. observations. Then the mean and covariance are the same for every observation, which is convenient for parameter estimation. With time series the i.i.d. assumption is problematic, but it is unnecessary. If the series ξ_t and ε_t are serially dependent, but ξ_t and ε_t are uncorrelated (at t) with zero means and constant covariances Γ and Ψ, then the mean and covariance of y_t are $\mu_y \equiv \alpha$ and $\Sigma_y \equiv B\Gamma B' + \Psi$, respectively. Under some regularity conditions the sample mean and covariance of y will be consistent estimators of μ_y and Σ_y, and therefore the usual estimators of the parameters (such as ML) are consistent. This principle is now demonstated under considerably weaker assumptions.

A slightly more general variant of (1) is used:

$$y_t = \alpha_t + B\xi_t + \varepsilon_t, \tag{2}$$

where α_t is a possibly time-varying intercept vector, but loadings are assumed time-invariant. Many time series integrate of order 1 so the variances of the indicators increase with time. This violates assumptions for standard estimators where parameters are constant and moments converge in probability to finite limits (see, e.g., Wansbeek and Meijer, 2000, p. 234).

Often y_t integrates but has a stationary first difference. Thus differencing is a common practice in time series analysis and the consequences of differencing (2) are examined. Below it is shown that assuming a stationary differenced series is stronger than necessary and a weaker form of boundedness suffices. Defining D as the difference operator (2) becomes

$$Dy_t \equiv y_t - y_{t-1} = (\alpha_t - \alpha_{t-1}) + B(\xi_t - \xi_{t-1}) + (\varepsilon_t - \varepsilon_{t-1})$$

or

$$Dy_t = \tau_t + B\,D\xi_t + D\varepsilon_t. \tag{3}$$

The latter is again an equation with a factor structure, and with the same loadings B. Thus a standard FA model can be estimated with the differenced data.

Following are sufficient conditions (assumptions) such that this leads to consistent estimators of relevant parameters. First, measurement model (2) and hence (3) is assumed. Second, it is assumed that $\tau_t = \tau$ is a constant vector in (3). In the application $\alpha_t = 0$ and therefore $\tau_t = 0$ for all t, but the theory is developed with the more general specification of non-zero but time-constant τ. Third, the following conditions are assumed:

1. $\kappa \equiv \operatorname{plim}_{T\to\infty} \sum_{t=1}^{T} D\xi_t/T$ exists and is finite.
2. $\operatorname{plim}_{T\to\infty} \sum_{t=1}^{T} D\varepsilon_t/T = 0$.
3. $\Phi \equiv \operatorname{plim}_{T\to\infty} \sum_{t=1}^{T} (D\xi_t - \kappa)(D\xi_t - \kappa)'/T$ exists and is finite and positive definite.
4. $\Omega \equiv \operatorname{plim}_{T\to\infty} \sum_{t=1}^{T} D\varepsilon_t\, D\varepsilon_t'/T$ exists and is finite and positive definite.
5. $\operatorname{plim}_{T\to\infty} \sum_{t=1}^{T} (D\xi_t - \kappa)\, D\varepsilon_t'/T = 0$.

Although unit roots in $D\xi_t$ and/or $D\varepsilon_t$ violate the assumptions, no other explicit assumptions are made about possible autocorrelation of the differenced data, and these assumptions allow considerable serial dependence in the variables.[1] Furthermore, it is not assumed that means and variances are constant over time, only that they are bounded in such a way that the required probability limits exist. This allows, for example, GARCH processes (Bollerslev, 1986).

Typically ξ_t is a random vector but alternatively it might be a series of given constants, in which case the measurement model is interpreted as a *functional model* not a *structural model* and "plim" has the same meaning as "lim" (Wansbeek and Meijer, 2000, pp. 11–12).

The conditions 2 and 5 are implied by the alternative condition $\mathrm{E}(D\varepsilon_t \mid D\xi_t) = 0$, combined with the finiteness of Φ and Ω. This is a substantively more meaningful assumption than 2 and 5 and therefore is assumed to be satisfied as well.

The sample mean and covariance of the differenced series Dy_t will be denoted by $\overline{Dy}$ and S_{Dy}, respectively. That is,

$$\overline{Dy} \equiv \frac{1}{T} \sum_{t=1}^{T} Dy_t$$

[1] If unit roots are present in the differenced series, the data can be differenced a second time, and the assumptions then apply to the twice differenced variables. The theory discussed here then also fully applies to the resulting analysis. This process can be repeated until no unit roots are present anymore.

and

$$S_{Dy} \equiv \frac{1}{T}\sum_{t=1}^{T}(Dy_t - \overline{Dy})(Dy_t - \overline{Dy})'.$$

From the stated assumptions, it follows that

$$\operatorname*{plim}_{T\to\infty} \overline{Dy} = \mu \equiv \tau + B\kappa \tag{4}$$

and

$$\operatorname*{plim}_{T\to\infty} S_{Dy} = \Sigma \equiv B\Phi B' + \Omega. \tag{5}$$

Conventional FA estimators (such as ML) use the sample covariance to estimate the loadings B, the factor covariance Φ, and the error covariance Ω. From (5) it follows that these estimators must also be consistent when S_{Dy} is used as the sample covariance. Neither normality nor serial independence are required for this result. However, just as in standard FA, consistency is only obtained if B, Φ, and Ω are identified from this equation (i.e., they are uniquely determined if Σ is known). Therefore it is assumed that this is the case. In the example in section 5, as in most applications, Ω is assumed to be diagonal. Then, if the Ledermann bound

$$(M-k)^2 \geq M + k$$

is satisfied, Ω is generally identified (Wansbeek and Meijer, 2000, pp. 169–170). As in standard FA, the parameter matrices B and Φ are uniquely defined either by imposing restrictions on their elements or by choosing a rotation method (see, e.g., Browne, 2001; Loehlin, 1987, chap. 6).

Given estimators $\hat{B}$, $\hat{\Phi}$, and $\hat{\Omega}$, estimators for τ and/or κ can be obtained from (4). The number of sample means in this equation is smaller than the number of parameters and therefore some restrictions must be imposed. In a typical FA model, the intercepts are free parameters, so that the means of the factors can be arbitrarily but conveniently restricted to zero, giving the restriction $\kappa = 0$ and estimator $\hat{\tau} = \overline{Dy}$. This illustrates why the means are usually neglected in FA applications. When $\tau = 0$ and κ is not zero, a natural and consistent estimator of κ is the GLS estimator

$$\hat{\kappa} = (\hat{B}'\hat{\Omega}^{-1}\hat{B})^{-1}\hat{B}'\hat{\Omega}^{-1}\overline{Dy}.$$

It is also possible to estimate all parameters jointly from the mean and covariance structure, i.e., use (4) and (5) jointly. Some experimentation with this did not lead to

improved estimators and attention is restricted to a standard covariance-based estimator of free parameters in B, Φ, and Ω. In particular, the maximum likelihood estimator found by minimizing

$$L \equiv \log\det\Sigma + \mathrm{tr}(\Sigma^{-1}S_{Dy}) \tag{6}$$

as a function of the parameters is used. Here, although not made explicit in the notation, Σ is a function of the parameters, as given in (5). Resulting consistent estimators will not be full maximum likelihood, but quasi maximum likelihood in the sense of White (1982). This is because the data are typically not normally distributed, may be serially dependent, and (4) may give additional information on the parameters (e.g., if $\tau = 0$), which is unused in the estimation.

Under weak assumptions, the central limit theorem implies that the elements of the sample covariance S_{Dy} are jointly asymptotically normally distributed. Let s_{Dy} be the vector consisting of all unique (nonduplicated) elements of S_{Dy}, and let σ_0 be its probability limit. Then

$$\sqrt{T}(s_{Dy} - \sigma_0) \xrightarrow{d} \mathcal{N}(0, \Upsilon_0) \tag{7}$$

for some finite positive definite matrix Υ_0. Υ_0 can be estimated consistently by a heteroskedasticity and autocorrelation consistent (HAC) covariance estimator, such as the Newey-West estimator (Newey and West, 1987). See Andrews (1991) and Wansbeek and Meijer (2000, pp. 249–252) for a discussion of HAC estimators and De Jong and Davidson (2000) for a very general consistency result for HAC estimators.

Stack B, Φ, and Ω in the parameter vector θ and denote the population value θ_0. The estimator $\hat{\theta}$ of θ is a function of s_{Dy}. Combining the implicit function theorem and the delta method with (7), gives

$$\sqrt{T}(\hat{\theta} - \theta_0) \xrightarrow{d} \mathcal{N}(0, J_0\Upsilon_0 J_0'),$$

where $J_0 \equiv \mathrm{plim}_{T\to\infty} \partial\hat{\theta}/\partial s_{Dy}'$. Formulas for $\partial\hat{\theta}/\partial s_{Dy}'$ were given by Shapiro (1983) for the case in which identification is obtained by explicit restrictions on the parameters. Archer and Jennrich (1973) and Jennrich (1973) derived formulas for the case in which a rotation method is used to obtain uniquely defined parameters. Standard errors of the parameter estimators are now straightforwardly obtained and Wald and LM tests can be routinely applied if desired.

3 Predicting factor scores

In many cases one is not only interested in the model parameters, such as B, but also, or even primarily, in the realized values of the factors, the *factor scores*. The factors are

unobserved, i.e., they are latent variables, and their values generally cannot be estimated consistently. They can, however, be "predicted." This is prediction for the same time period and should not be confused with prediction of the future (i.e., forecasting). In the economics and time series literature the term *estimation* is often used with indirectly measured latent variables, which often correspond to the states of state space models. Technically these are not estimates, since they do not converge to the values of the latent variables with increasing time, but the terminology does have the advantage of more easily distinguishing prediction of the future from prediction of the present and past.

Factor prediction error is due to two sources: (1) estimation error in the parameters, and (2) inability to obtain perfect predictions even if parameters are known. Parameter estimation error is of order $O_p(T^{-1/2})$ and in large samples this error should be small. The second source of error does not diminish with sample size, because, without assumptions about dynamic dependencies of the observations, y_s for $s \neq t$ do not provide information about ξ_t. Therefore, all usable information about ξ_t is contained in y_t and this information does not increase with sample size. It follows that this error is $O_p(1)$. Consequently, in large samples the prediction error is dominated by the second source of error. Discussion in the literature focuses on asymptotics, so parameters are assumed known.

Compared to standard factor score prediction literature, (2) is different in that means and intercepts can be nonzero. Standard formulas must be slightly adapted to accommodate this. Anderson and Rubin (1956) provided formulas for nonzero intercepts, but still assumed zero means of the factors. The extension to nonzero means is given below. As in standard factor score prediction, attention is restricted to linear predictors. Nonlinear predictors have been proposed in Meijer and Wansbeek (1999).

The easiest way to find suitable factor score predictors when intercepts and means are possibly nonzero is to transform the model such that the means and intercepts are zero, then apply the standard predictors to the transformed model and transform back. Starting from the model (2) with possibly nonzero means and intercepts, the transformed model is

$$y_t - \alpha_t - B\gamma_t = B(\xi_t - \gamma_t) + \varepsilon_t, \tag{8}$$

where $\gamma_t \equiv \mathrm{E}(\xi_t)$ and using the assumption $\mathrm{E}(\varepsilon_t) = 0$. When the parameters are assumed known, as discussed above, the left-hand side of (8) is a vector of observed indicators with mean zero, and $\xi_t - \gamma_t$ in the right-hand side is a vector of factors with mean zero.

The two most frequently used factor score predictors for a model with zero means and intercepts are the *regression predictor* and the *Bartlett predictor*. See, e.g., Wansbeek and Meijer (2000, pp. 164–166) for their derivation.

Applied to (8), the regression predictor for the zero-mean factor $(\xi_t - \gamma_t)$ is

$\Gamma_t B' \Sigma_{yt}^{-1}(y_t - \alpha_t - B\gamma_t)$, so that the resulting regression predictor for ξ_t is

$$\hat{\xi}_t^R = \gamma_t + \Gamma_t B' \Sigma_{yt}^{-1}(y_t - \alpha_t - B\gamma_t),$$

where $\Gamma_t \equiv \text{Cov}(\xi_t)$ and $\Sigma_{yt} \equiv \text{Cov}(y_t)$. Similarly, the Bartlett predictor for $(\xi_t - \gamma_t)$ is $(B'\Psi_t^{-1}B)^{-1}B'\Psi_t^{-1}(y_t - \alpha_t - B\gamma_t)$, so that the resulting Bartlett predictor for ξ_t is

$$\hat{\xi}_t^B = (B'\Psi_t^{-1}B)^{-1}B'\Psi_t^{-1}(y_t - \alpha_t),$$

where $\Psi_t \equiv \text{Cov}(\varepsilon_t)$. The (unconditional) means and covariance matrices of the predictors are

$$\begin{aligned}
\text{E}(\hat{\xi}_t^R) &= \Gamma_t B' \Sigma_{yt}^{-1}(\alpha_t + B\gamma_t) - \Gamma_t B' \Sigma_{yt}^{-1}\alpha_t - (\Gamma_t B' \Sigma_{yt}^{-1} B - I_k)\gamma_t \\
&= \gamma_t \\
\text{Cov}(\hat{\xi}_t^R) &= \Gamma_t B' \Sigma_{yt}^{-1}(\Sigma_{yt})\Sigma_{yt}^{-1} B\Gamma_t \\
&= \Gamma_t - \Lambda^{-1} \\
&< \Gamma_t
\end{aligned}$$

and

$$\begin{aligned}
\text{E}(\hat{\xi}_t^B) &= (B'\Psi_t^{-1}B)^{-1}B'\Psi_t^{-1}(\alpha_t + B\gamma_t) - (B'\Psi_t^{-1}B)^{-1}B'\Psi_t^{-1}\alpha_t \\
&= \gamma_t \\
\text{Cov}(\hat{\xi}_t^B) &= (B'\Psi_t^{-1}B)^{-1}B'\Psi_t^{-1}(\Sigma_{yt})\Psi_t^{-1}B(B'\Psi_t^{-1}B)^{-1} \\
&= \Gamma_t + (B'\Psi_t^{-1}B)^{-1} \\
&> \Gamma_t,
\end{aligned}$$

where $\Lambda \equiv \Gamma_t^{-1} + B'\Psi_t^{-1}B$ and $\Lambda^{-1} = \Gamma_t - \Gamma_t B' \Sigma_{yt}^{-1} B\Gamma_t$, cf. Wansbeek and Meijer (2000, pp. 164–165), and matrix inequalities are in the sense of Löwner (1934), i.e., $A < B$ means that $B - A$ is positive definite. Thus, the means of the predictors are the same as the mean of ξ_t, but the variances are different.[2]

Both predictors require knowledge of α_t, the intercept, and Ψ_t, the covariance of ε_t. In addition to this, the regression predictor requires knowledge of γ_t and Γ_t, the mean and covariance of ξ_t, whereas the Bartlett predictor does not. This is an important difference, because these are generally not known and assumptions about ξ_t are to be minimized. Moreover, with an integrated series, γ_t and Γ_t increase with t and any

[2] *Covariance preserving* predictors are not used here but can be developed if it is desirable that the predictors have the same covariance as the factors. See Ten Berge et al. (1999).

assumptions become more problematic. In the functional model view mentioned earlier, it is even more questionable whether the MSE optimality of the regression predictor is meaningful. In this case, the Bartlett predictor is a perfectly natural estimator of an unknown vector.

Knowledge of Ψ_t is still needed for the Bartlett predictor. However, it is substantively and interpretationally convenient to assume $\Psi_t \equiv \mathrm{Cov}(\varepsilon_t) = \Psi$ is a time-invariant diagonal matrix and ε_t and ε_s are independent for $t \neq s$. This assumption about ε_t implies that the covariance of $D\varepsilon_t$ is $\Omega = 2\Psi$. It now follows that Ψ_t may be replaced by Ω in the definition of the Bartlett predictor. Because Ω is consistently estimated from the differenced data, it may be assumed known asymptotically. If the i.i.d. assumption about ε is not met, and any positive definite weight matrix W is inserted for Ψ_t^{-1} in the computation of the Bartlett predictor, this predictor is still unbiased, although not optimally efficient. This follows from standard GLS regression theory. Therefore, the Bartlett predictor with Ω^{-1} inserted for Ψ_t^{-1} generally makes sense and will have relatively good properties even in the non-i.i.d. case.

When it is assumed that $\alpha_t = \alpha = 0$, the regression predictor becomes

$$\hat{\xi}_t^R = \gamma_t + \Gamma_t B' \Sigma_{yt}^{-1}(y_t - B\gamma_t)$$

and the Bartlett predictor becomes

$$\hat{\xi}_t^B = (B'\Psi_t^{-1}B)^{-1}B'\Psi_t^{-1}y_t$$

or

$$\hat{\xi}_t^B = (B'\Omega^{-1}B)^{-1}B'\Omega^{-1}y_t.$$

The latter formula does not contain any possibly time-varying parameters. This further enhances the advantage of the Bartlett predictor over the regression predictor, because the Bartlett predictor can be computed by using just the estimation results from the factor analysis of the differenced series, in particular B and Ω. From this it is clear that we have a strong preference for the Bartlett predictor in a time-series context.

4 Simulation study

The precision of estimators and factor score predictors is assessed with a small simulation study. To give a realistic view, the simulation used true factors $\tilde{\xi}_t$ and model parameters based on those estimated from a real data set. The real estimation is described in section 5. There are $M = 6$ indicators and $k = 2$ factors, and data for $T = 215$ consecutive months. These true factors and parameters were used to generate

new samples. In this way, the new samples should be similar to data encountered in practice. Simulated data was generated for 100 replications (samples).

Parameters were estimated with standard ML applied to the differenced data, as discussed in section 2. The direct oblimin (quartimin) rotation method with Kaiser normalization was used (Wansbeek and Meijer, 2000, pp. 168–169). This is a common rotation method if the factors are not assumed (contemporaneously) uncorrelated. It usually gives clearly interpretable results. The factor score predictor was the Bartlett predictor, as argued in the previous section. However, as shown there, the covariance of these predictors is larger than the covariance of the actual factors. Therefore, the predicted scores were linearly transformed such that the (sample) covariance of their first differences was exactly equal to the estimated covariance $\hat{\Phi}$ of the first differences of the factors. Otherwise, the true parameter values in the simulation would not equal the estimates from the real data.

All replications used the same values of the factors and thus implications of the simulation are conditional upon these. They are depicted by the solid lines in figure 1, whereas their first differences are depicted by the solid lines in figure 2. These figures suggest that the factors are trending and have a unit root and that their first differences have some negative serial correlation.

New values $\varepsilon_t^{(r)}$ of the errors, where r is the replication number, were drawn from a normal distribution with mean zero and covariance $\hat{\Psi} = \frac{1}{2}\hat{\Omega}$ given by the estimated covariance from the original data. New sample data were subsequently obtained from

$$y_t^{(r)} = \hat{B}\tilde{\xi}_t + \varepsilon_t^{(r)},$$

where $\hat{B}$ is the estimated loadings from the original data.

Estimation and factor score prediction from the simulated data was done as estimation and factor score prediction for the original data, but the transformation of the factor score predictors was omitted, because this would make the predictors biased. (This is also omitted in the application described in the next section.)

Computation was done with the software R (R Development Core Team, 2004), using the gpa functions of Bernaards and Jennrich (in press) for the rotations.[3]

The first replication

The first generated sample illustrates what can be expected in an empirical situation. This sample was used to estimate the parameters B and Ω and calculate predicted factor scores. Figure 1 depicts true factor scores (solid lines) and Bartlett predictions (heavy

[3]We would like to thank Coen A. Bernaards and Robert I. Jennrich for extensive discussions about their code.

dashed lines), using sample estimates of the parameters. Other lines will be discussed further below.

There appears to be a large and systematic bias in the predicted scores. To analyze this, consider the Bartlett prediction error:

$$\begin{aligned}\hat{\xi}_t - \xi_t &= (\hat{B}'\hat{\Omega}^{-1}\hat{B})^{-1}\hat{B}'\hat{\Omega}^{-1}y_t - \xi_t \\ &= (\hat{B}'\hat{\Omega}^{-1}\hat{B})^{-1}\hat{B}'\hat{\Omega}^{-1}(B\xi_t + \varepsilon_t) - \xi_t \\ &= -\hat{L}'(\hat{B} - B)\xi_t + \hat{L}'\varepsilon_t, \end{aligned} \tag{9}$$

where $\hat{L}' = (\hat{B}'\hat{\Omega}^{-1}\hat{B})^{-1}\hat{B}'\hat{\Omega}^{-1}$. The first term on the right-hand side of (9) is $O_p(T^{-1/2})$, and zero if $\hat{B} = B$. The second term is $O_p(1)$ and does not converge to zero with increasing sample size unless $L'\varepsilon_t = 0$, where $L' = (B'\Omega^{-1}B)^{-1}B'\Omega^{-1}$, but this can be neglected as it has zero probability.

Since ε_t in the simulation are i.i.d. with mean zero, the second term in (9) is nonsystematic with a zero mean. Therefore, the systematic error cannot be due to this term and must come from the first one. There is a thin dashed line in figure 1 which is virtually indistinguishable from the true value. It plots the predicted scores that would result if the true parameter values B and Ω are used instead of their sample estimates in the computation of the predicted factor scores. Predictions are extremely good and there is no noticeable bias. This illustrates that the prediction errors are largely due to estimation errors of the parameters. Apparently, the sample size is (much) too small for the asymptotic analysis of the prediction error. Asymptotically, the estimation error is a negligible part of the prediction error, because the former is $O_p(T^{-1/2})$ and the latter is $O_p(1)$, but here the estimation error explains almost all of the prediction error.

The systematic nature of the prediction errors is explained by the positivity of ξ_t and the fact that $\hat{L}'(\hat{B} - B)$ is a constant matrix in a given sample. Thus errors tend to have the same sign for all t. Moreover, if the ξ_t are serially dependent, the prediction errors will also be serially dependent and prediction errors of consecutive time points tend to be approximately the same. Finally, because ξ_t is generally increasing with t, it follows from (9) that the prediction errors become larger for more recent time points.

The effect is much smaller when comparing first differences of the predicted factor scores with the first differences of the true factor scores, because the first differences can be negative as well and they are considerably less serially dependent. This is illustrated in figure 2.

Observable bias and bias correction

From the factor score predictions $\hat{\xi}_t$, predicted values of indicators can be computed by

$$\hat{y}_t = \hat{B}\hat{\xi}_t, \tag{10}$$

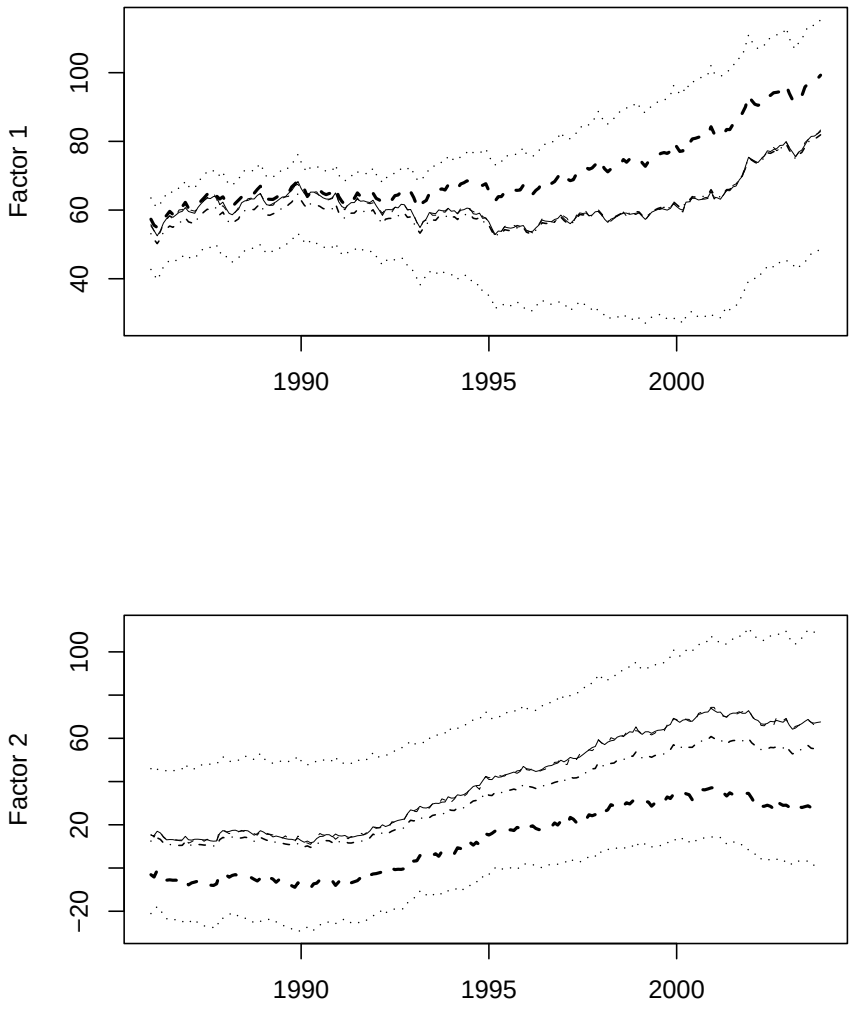

Figure 1: True factor scores (solid lines); Bartlett factor score predictions (using sample estimates of the parameters; heavy dashed lines) in the first replication; Bartlett factor score predictions (using true values of the parameters; thin dashed lines) in the first replication; Bartlett factor score predictions (means and means ± 2 s.d.), using 100 replications (dot-dashed and dotted lines).

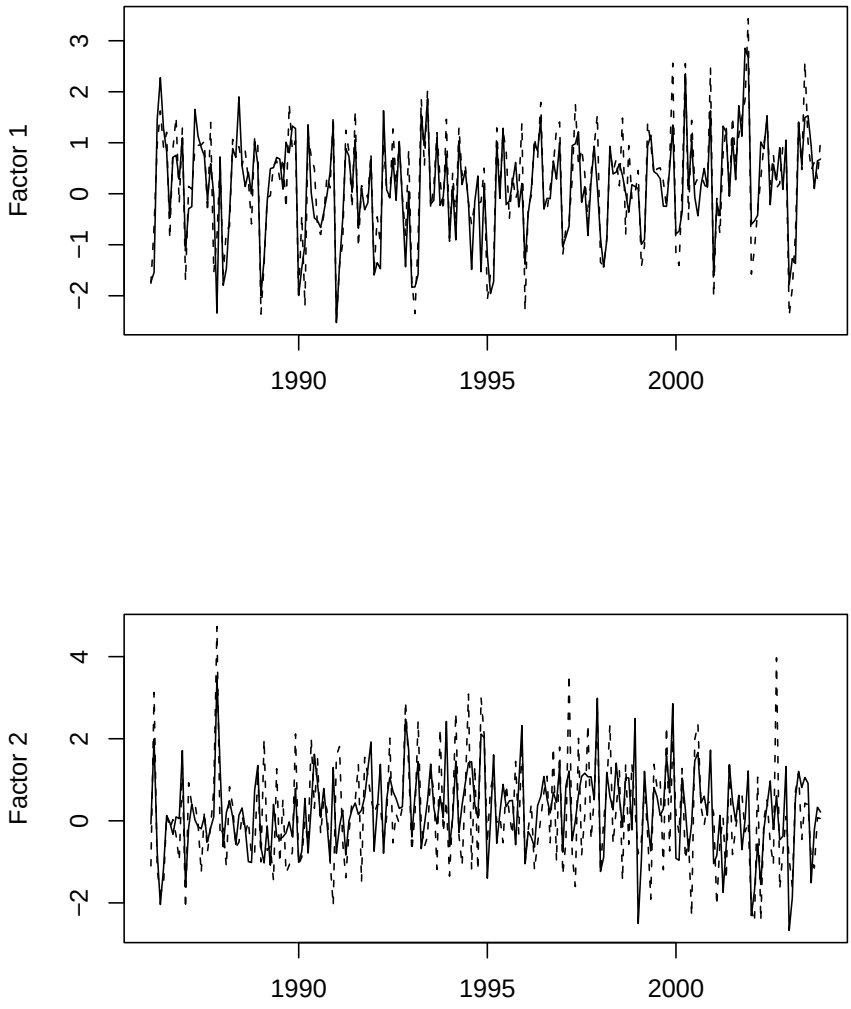

Figure 2: First differences of true factor scores (solid lines) vs. first differences of Bartlett factor score predictions (dashed lines) in the first replication.

and these can be compared with observed values of the indicators. The TSFA model does not assume that the errors ε_t should be small, it only assumes that they are uncorrelated (at the same time point) and have mean zero. But this comparison gives an indication about the fit of the model, in addition to likelihood ratio statistics and other fit measures. For the first replication, the predicted values of indicators 1–3 are plotted with the observed values in figure 3 and the predicted values of indicators 4–6 are plotted with the observed values in figure 4. Here, it can be seen that the results for the first three indicators are representative of the other indicators, so in the sequel, the figures of indicators 4–6 are omitted for space considerations.

The systematic factor score prediction errors largely carry over to prediction errors in observed indicators. The latter, being observed, suggests that some kind of bias correction might be performed, so that indicator prediction errors vanish. This might then reduce systematic errors in predictors, but this is a false hope. The differences between the measured values of the indicators and their predicted values are

$$y_t - \hat{y}_t = y_t - \hat{B}\hat{\xi}_t.$$

It seems natural to define the optimal predictors as the ones that minimize a generalized least squares function

$$F \equiv (y_t - \hat{y}_t)' W (y_t - \hat{y}_t),$$

where W is a weight matrix. From standard regression theory the optimal predictor is

$$\hat{\xi}_t = (\hat{B}' W \hat{B})^{-1} \hat{B}' W y_t.$$

Analogous to GLS regression theory, it is natural to choose $W = \hat{\Omega}^{-1}$ here. Consequently, the optimal factor score predictor is the Bartlett predictor. Given that the analysis did not involve the other observations, this predictor is optimal for each observation separately and therefore for all observations jointly as well. Although it looks like there is a systematic bias in the predictors and this is carried over to the predicted indicators, it is not possible to obtain a better (bias-corrected) factor score predictor that reduces the bias in the observed indicators.

More replications

Moving on to consider all 100 replications, table 1 shows the bias and variability of estimated loadings. With the possible exception of the bottom-left element there are no substantial biases. However, there is a large variability in the estimates.

Figure 1 plots the means (across replications) of the predictions with the dot-dashed line and the means plus and minus 2 times the standard deviations with the dotted lines. The latter two give an impression of the variability of the predictors around the means.

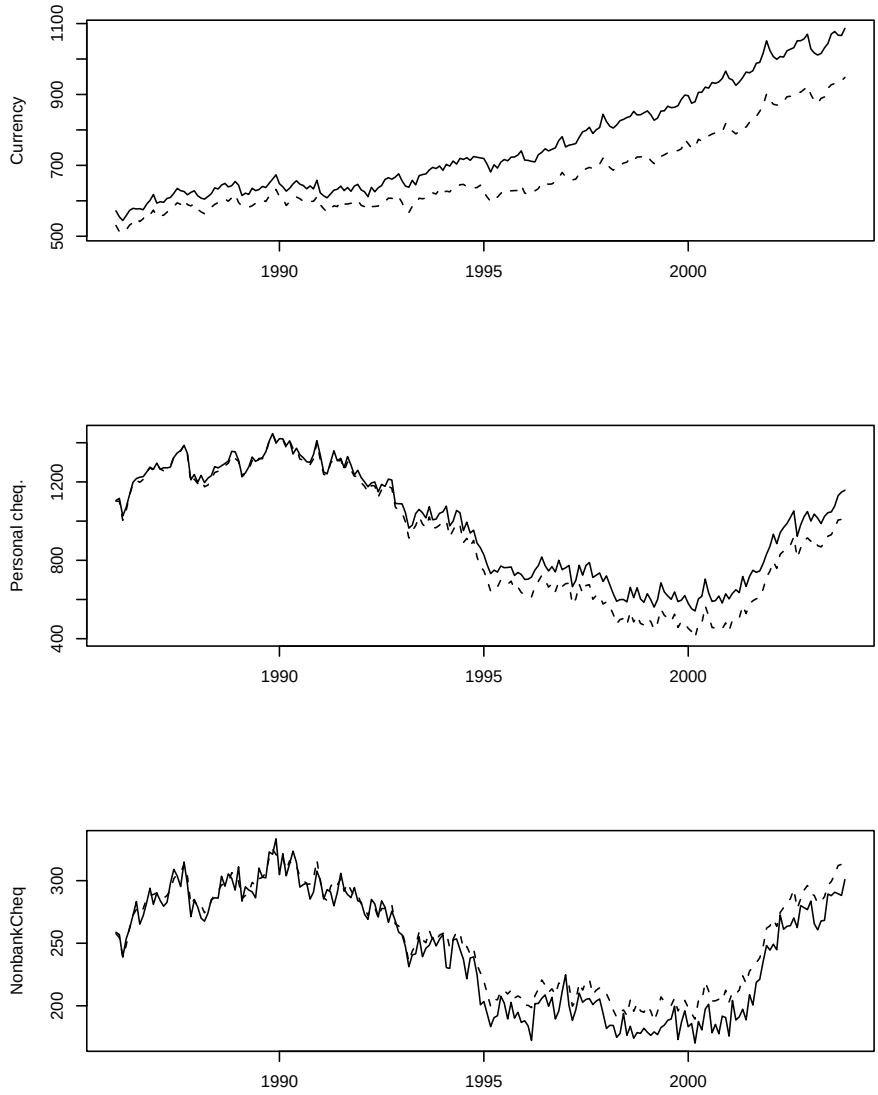

Figure 3: Observed (solid lines) vs. predicted values (dashed lines) of indicators 1–3 in the first replication.

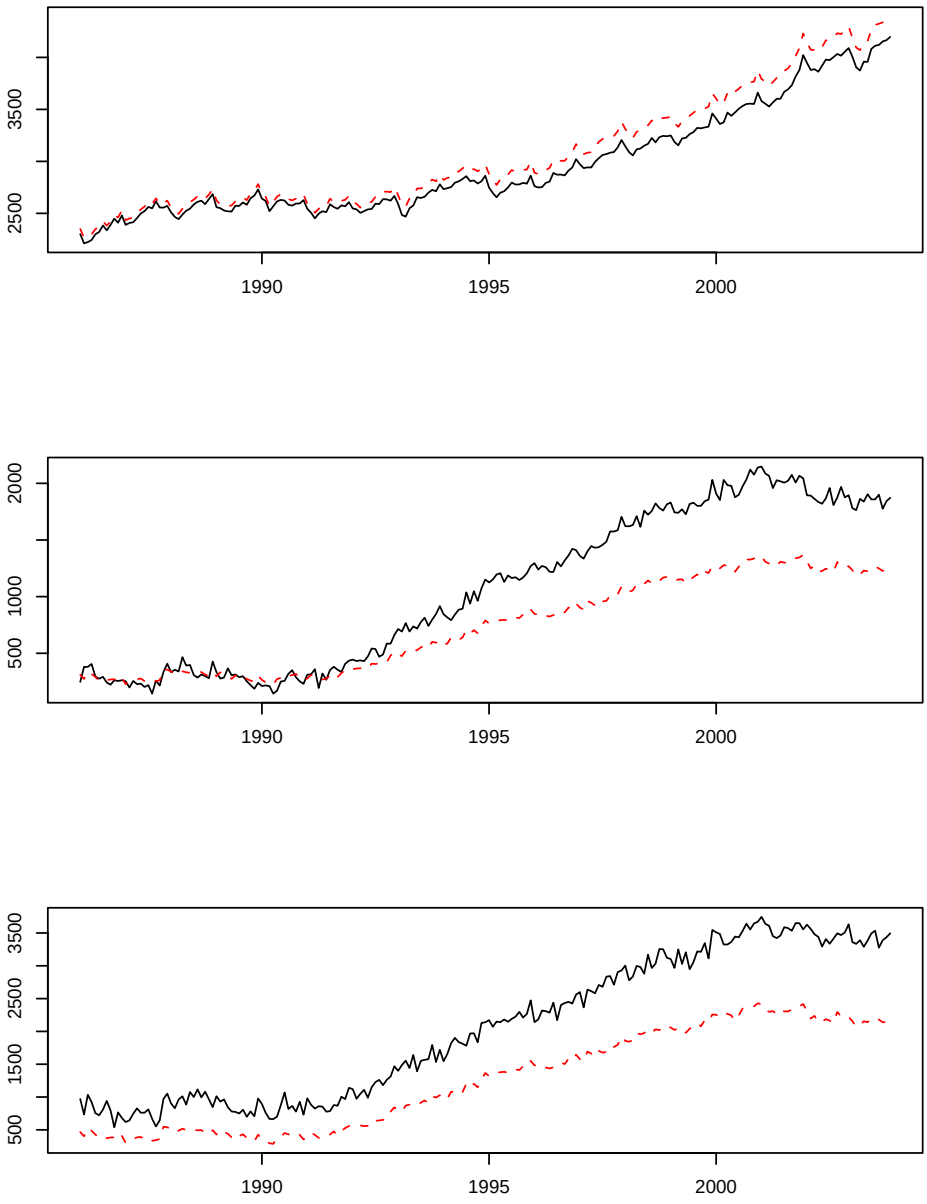

Figure 4: Observed (solid lines) vs. predicted values (using sample estimates of the parameters; dashed lines) of indicators 4–6 in the first replication.

Table 1: Bias and variability in the factor loadings estimates.

Indicator	True values		Bias		Standard deviation	
	Factor 1	Factor 2	Factor 1	Factor 2	Factor 1	Factor 2
1	8.8	5.2	-0.4	-0.7	2.1	2.9
2	23.8	-12.6	-2.5	0.4	5.9	8.9
3	5.2	-2.0	-0.4	-0.1	1.5	1.9
4	36.8	16.9	-1.3	-2.5	6.9	11.0
5	-2.8	31.0	1.0	0.4	9.6	10.5
6	2.6	47.6	2.5	-0.4	12.4	17.4

The figure shows that the systematic prediction errors encountered before are not due to a bias, because there is little sign of bias across replications (with the possible exception of a small bias in the second factor in later time points). This corroborates the earlier analysis. However, there is considerable variability around the means which, from the earlier analysis, is mainly due to the large variability in the parameter estimates (especially the factor loadings).

5 Application to money data

In this section the techniques developed in sections 2 and 3 are applied to the Canadian money data. This updates the application described in Gilbert and Pichette (2003) using estimation techniques developed above. Also, a different rotation criterion is used, but rotation is not the primary focus of the current work.

As previously mentioned, there are problems with current monetary aggregates, and predictors based on TSFA may better measure the concepts of interest. The intention is to eventually replace traditional monetary aggregates with index measures using TSFA. That is, aggregates will be replaced with factors which are the latent variables explaining shifts in the financial assets of the population. These factors are predicted using the component data from the monetary aggregates as the indicators.

The indicators are organized into six categories. Data is measured in a larger number of categories, but shifts between indicators for reasons other than fundamental macro-economic behaviour will interfere with extracting the factors of interest. Thus it is necessary to combine asset categories among which there have been important shifts for other reasons. For example, Canadian Savings Bonds are no longer widely used and people have shifted to mutual funds. Instruments typically thought to be savings are grouped together and called investment. The six indicators are *currency*,

personal chequing deposits, *non-bank chequing deposits*, *non-personal demand and notice deposits*, *non-personal term deposits*, and *investment*. Non-personal accounts mostly belong to businesses. In Canada, trust companies, credit unions, and some other financial institutions are not grouped with banks, so the term "non-bank" refers to deposits at these institutions. Considerably more detail about the data is provided in Gilbert and Pichette (2003) and in Kottaras (2003). To remove factors that are not the main interest, the data is per capita and measured in real Canadian dollars. Data is for the 215 months from January 1986 to November 2003.

Some indicators show no seasonality, while others do. For example, currency has marked seasonal peaks at Christmas and in September. These different patterns in the indicators may reflect differences in the factors. Thus seasonality may help distinguish the factors of interest, and therefore the data is not seasonally adjusted. For economic modeling the predicted factor scores could be seasonally adjusted.

Eigenvalues of the sample correlation matrix of differenced indicators give a rough idea of the number of factors to consider. These are 2.08, 1.39, 0.85, 0.69, 0.65, and 0.33. A conventional rule of thumb is that the number of factors should be equal to the number of eigenvalues larger than 1, suggesting at least 2 factors. One factor representing *transactions money* (assets intended for imminent transactions), and another representing *savings* (assets intended for longer term investment) was anticipated. However, the third eigenvalue is close to 1 suggesting a possible third factor. For example, corporate finances may be arranged differently than personal finances. Some preliminary experimentation with a third factor seems to suggest this. However, extracting three factors from six indicators causes statistical difficulties (the Ledermann bound is satisfied as an equality). Therefore, this possibility is left for later study with additional indicators.

A direct oblimin (quartimin) rotation was used to obtain a unique solution. This is arguably the most common non-orthogonal rotation method. The oblimin objective is to give loadings that weight heavily on one factor or the other. This may be appropriate for currency and investment, at the two ends of the scale, but is probably not appropriate for some deposit types in between. Modern deposit accounts can both pay good interest rates and allow the convenience of point-of-sale payments, so they may be used for both savings and transactions. Therefore, rotation criteria will be examined more carefully in future work.

The model chi-square (likelihood ratio statistic) is 3.19 with 4 degrees of freedom. The comparative fit index (CFI, a pseudo-R^2) is 1.0, because the model chi-square is less than its degrees of freedom. See Wansbeek and Meijer (2000, chap. 10) for a discussion of these and other fit measures. Although the assumptions on which these statistics are based (normally distributed i.i.d. data) are not met, and thus standard p-values are

Table 2: Estimation results for differenced data.

Indicator	Unstandardized loadings		Standardized loadings		Communality
	Factor 1	Factor 2	Factor 1	Factor 2	
Currency	8.84	5.20	0.66	0.39	0.59
Personal cheq.	23.82	-12.57	0.54	-0.28	0.37
NonbankCheq	5.18	-1.97	0.48	-0.18	0.26
N-P demand and notice	36.78	16.94	0.77	0.35	0.72
N-P term	-2.84	31.02	-0.04	0.44	0.20
Investment	2.60	47.63	0.02	0.40	0.16

incorrect, these findings indicate that the model fits very well.

The estimation results are shown in table 2. The two left columns are the estimated (unstandardized) loadings. These are the same as the loadings used previously to generate simulated data, because of the way that those loadings were constructed. The next two columns are the loadings of the standardized solution (scaled so that each series in the factors and data has variance 1.0). These are typically easier to interpret, because a larger (in absolute size) loading indicates a stronger relationship. Loadings of the standardized solution are typically between -1 and 1. This is necessarily the case for orthogonal solutions, because then the loadings are the correlations between the indicators and the factors. In principle, absolute loadings could be larger than 1 for oblique rotations such as oblimin, but such large loadings are rare. The estimated correlation between the two (differenced) factors is 0.0095, so the solution is nearly orthogonal. The last column gives the communalities. The communality of an indicator is its common variance (i.e., the variance explained by the factors) expressed as a fraction of its total variance. Thus, it is the R^2 of the regression of the indicator on the factors.

The communalities of most indicators are quite low compared to the communalities commonly encountered in cross-sectional analyses (e.g., De Haan et al., 2003). This is due to the differencing of the data, which tends to reduce the variability of the factors considerably and tends to *increase* the variability of the random errors. The combined effect of this is a reduction in the communality, which is the relative contribution of the factors. This effect is similar to the effect commonly encountered in panel data regressions, where the R^2 is much smaller for models in first differences than it is for models in levels. Tentative communalities are computed for the undifferenced data under the assumptions that the random errors are i.i.d. and the relevant covariance of

the factors is the covariance of the factor scores that were used in the simulation. The resulting communalities are 0.998, 0.992, 0.979, 0.999, 0.996, and 0.995, respectively, which illustrates the dramatic effect of differencing. We do not regard the low communalities for the differenced data as problematic. Extremely high communalities for the undifferenced data may be due to truly common factors as well as a spurious time dependence. One advantage of the differencing is that the latter is also removed and thus has no detrimental effect on parameter estimation.

The last two indicators only load on the second factor, which can therefore be interpreted as the savings factor, as anticipated. The relatively high loadings of the first few indicators on the first factor give some credence to its interpretation as transactions money. However, the moderately high loadings of the first and fourth indicators on the second factor, and the moderately high negative loadings of the second and third indicators on the second factor complicate the tentative interpretation of the factors. Even disregarding the difficulty interpreting negative loadings in this application, it would mean that currency is used for savings much more than usually thought. Another possibility would be that the oblimin rotation does not give the substantively best solution and another rotation would represent transactions money and savings money better. A more likely possibility is that there needs to be a third factor. For example, corporate financial behaviour may be very different from personal behavior. This conclusion is also suggested by the fact that non-personal demand and notice deposits also load heavily on the first factor.

Three factors is the Ledermann bound with six data series, which means that the covariance structure would be perfectly explained by the factors, regardless of the data. Therefore, it is difficult to assess the fit of the model with 3 factors. Furthermore, results may not be very meaningful due to overfitting. In future research this will be addressed by adding credit series.

Despite these problems, the Bartlett predictor will be illustrated. The data explained by two factors, using formula (10), which has the obvious shortcomings discussed earlier (but does not depend on the specific rotation method chosen), is shown in figure 5. This figure shows systematic discrepancies that are qualitatively similar to the corresponding results from the simulation study (figure 3). Furthermore, the seasonal pattern in currency is also clearly present in its prediction, whereas for investment (figure omitted) neither the observed indicator nor its prediction exhibits seasonality. The differenced version, which suggests that the factors capture movements fairly well, is shown in figure 6.

While the estimates are preliminary in several ways, it is still interesting to compare the results with other measures. The predicted factor scores are shown in figure 7, plotted against real per capita M1 and real per capita M2++. The predictors are scaled to

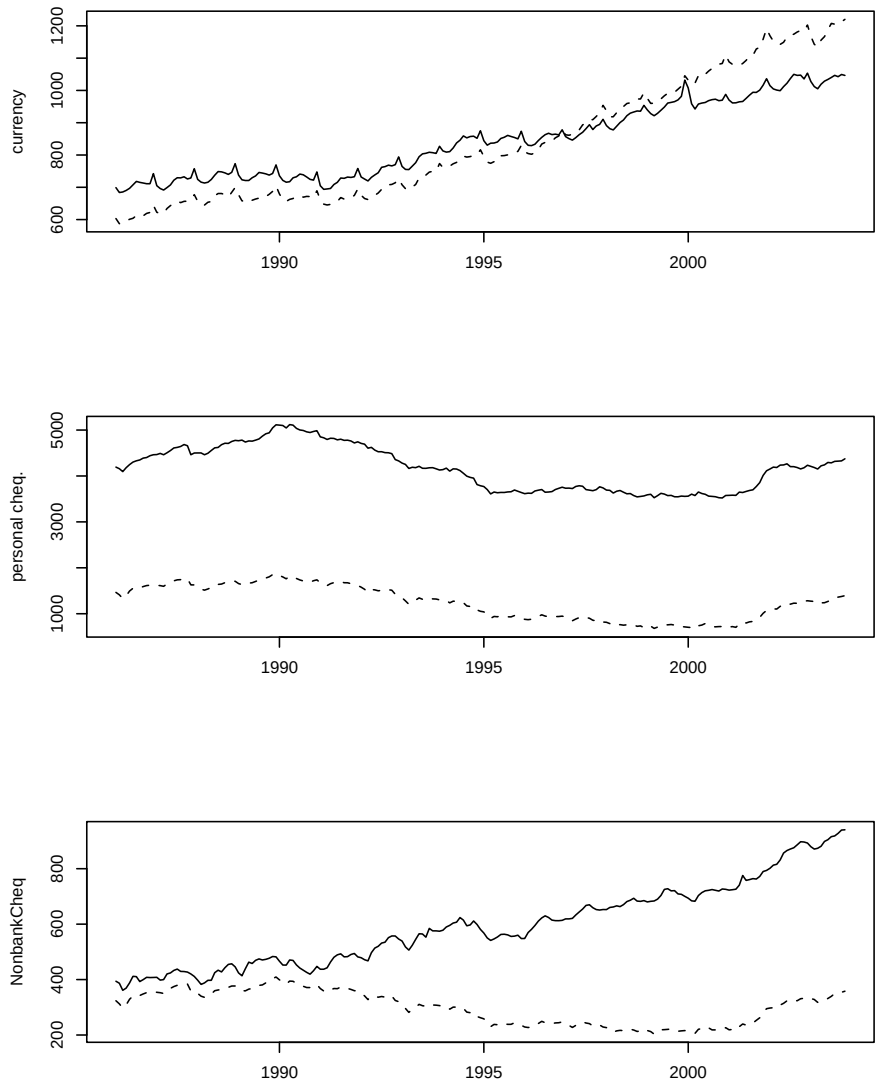

Figure 5: Explained money indicators 1–3 (solid lines: observed, dashed lines: predicted).

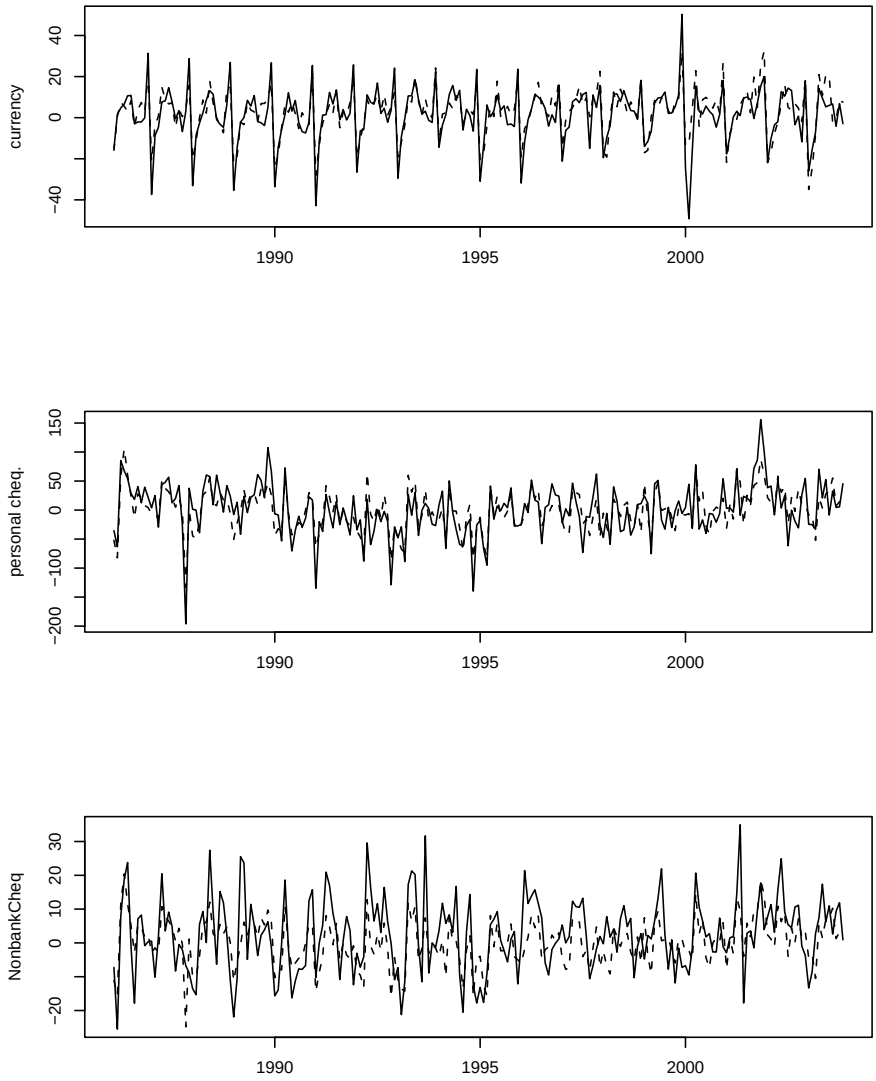

Figure 6: Explained money indicators 1–3, differenced (solid lines: observed, dashed lines: predicted).

have the same average values as M1 and M2++. The savings predictor has a much more dynamic path than M2++, and is also roughly consistent with what one might expect for savings given the Canadian business cycle. The first predictor has less pronounced growth than M1 in the second half of the sample. If one thinks that transactions money growth should be correlated with inflation then this less pronounced growth is certainly consistent with observed inflation. However, the drop of inflation in the late 1980's is not so clearly reflected.

The differenced predictors and differenced real per capita M1 and real per capita M2++ are shown in figure 8. The differenced first predictor looks surprisingly close to differenced M1, but this is more apparent than real, because the differenced first predictor is lower on average. One would expect a transaction money measure to be not too unlike M1 but, given that this whole research program was initiated because of dissatisfaction with M1 and other narrow aggregates, one might expect that it should not be too close either. Thus it is hard to pass judgement on this initial result. The second predictor shows considerably more volatility than M2++, and again it is difficult to immediately judge whether that is good or bad.

6 Sensitivity to sample period and size

This section provides an indication of the extent to which sample size and the selected sample period may influence results. By most accounts, the most recent data, starting in the mid-1990s, has been problematic because of financial innovations. Figures in this section shows the results estimated on different sub-samples: (1) January 1986–December 1989, (2) January 1986–December 1994, (3) January 1995–November 2003, and (4) January 2001–November 2003. Parameter estimators based on a (smaller) subsample are more variable than estimators based on the whole sample, so differences are expected.

Figures 9 and 10 show the predictions of the indicators and their first differences, according to formula (10). Although predictions for some periods appear to better than for others, this appears not to be systematic across indicators and is most likely due to the variability. So the level data generally has the same problems illustrated previously. The differenced data is relatively well explained on all subsamples. Most importantly, there is no convincing evidence that the predictions from the subsamples are noticably better than those from the complete sample, which gives some indication that the measurement model does not vary over time.

The predictors are shown in figure 11. The results from different subsamples are not extremely different from results for the complete sample, with the exception of the post-2001 subsample. This is the smallest subsample ($T = 35$), so is probably

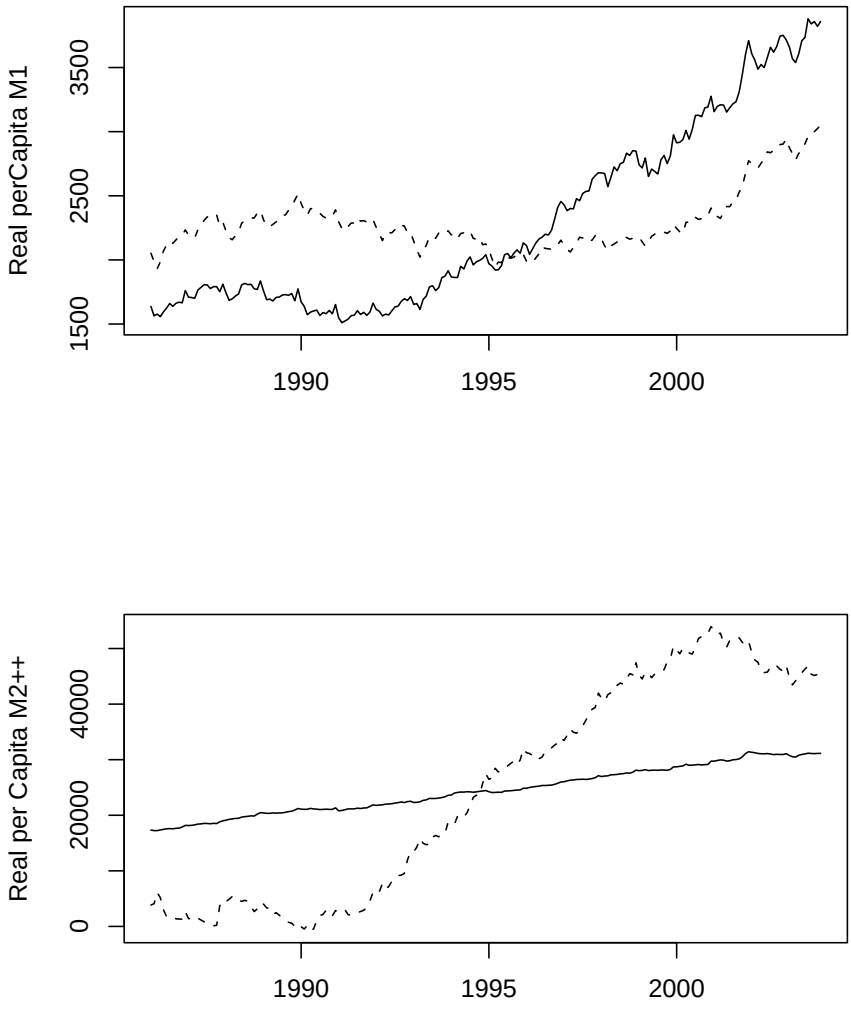

Figure 7: M1 and M2++ (solid) and scaled Bartlett predictors (dashed).

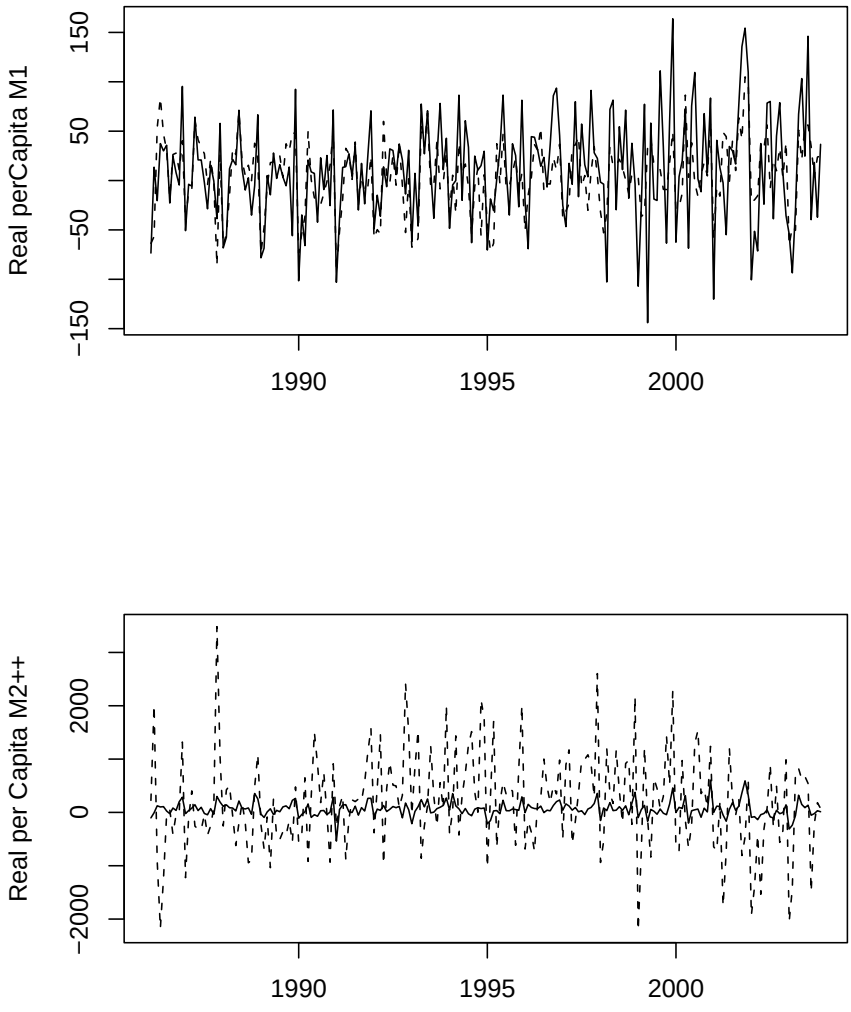

Figure 8: Differenced M1 and M2++ (solid) and scaled Bartlett predictors (dashed).

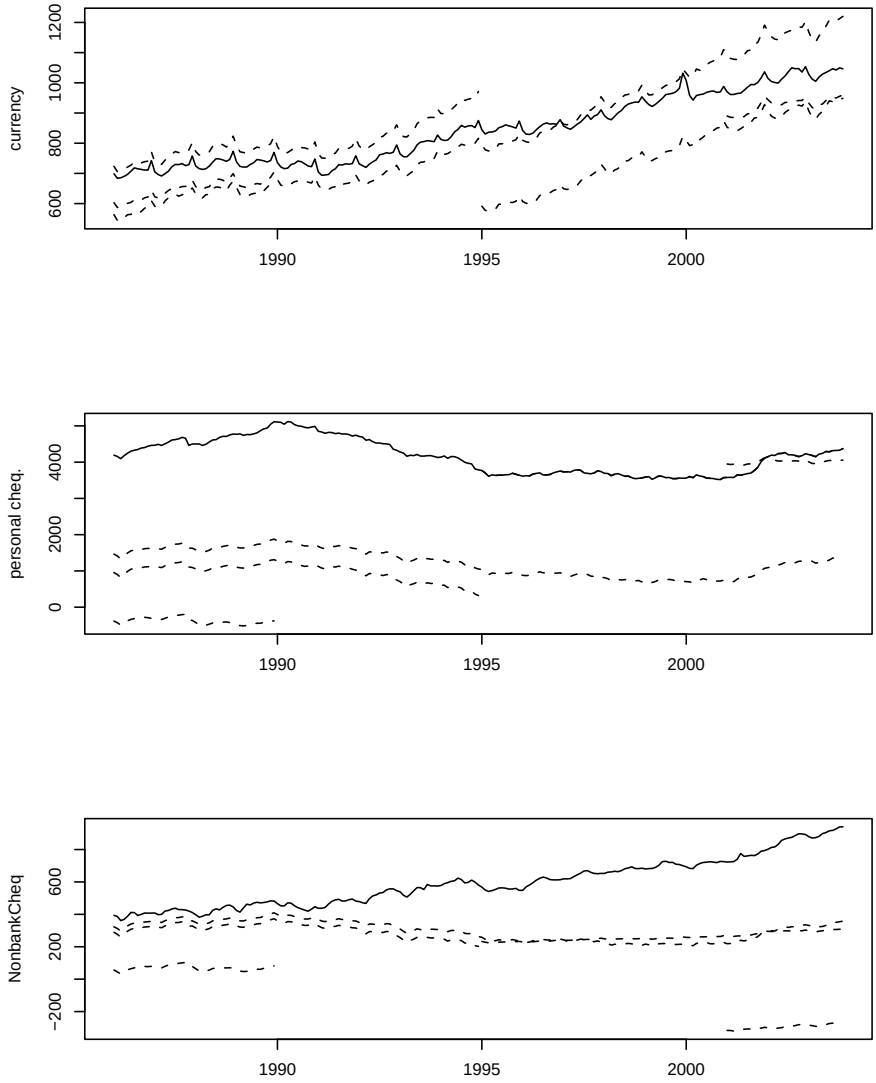

Figure 9: Explained money indicators 1–3, full sample and sub-samples (solid: observed, other: predicted using different subsamples).

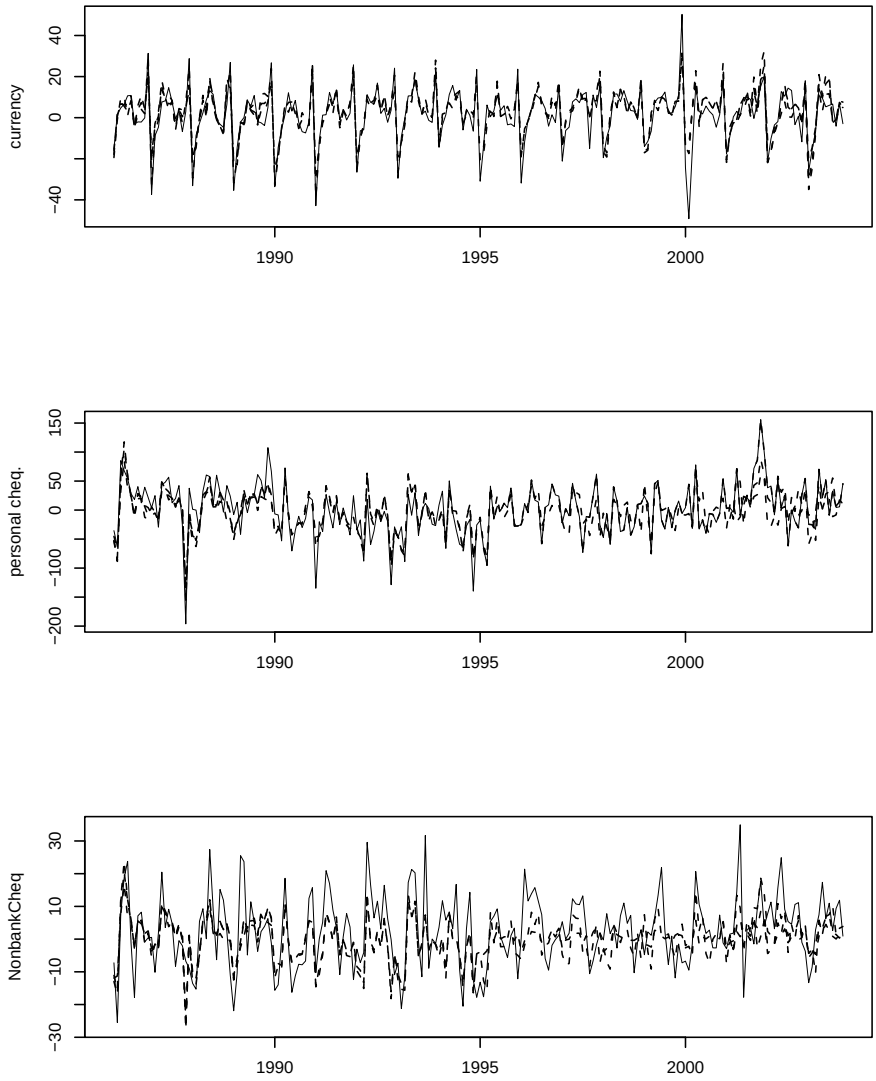

Figure 10: Explained money indicators 1–3, differenced, full sample and sub-samples (solid: observed, other: predicted using different subsamples).

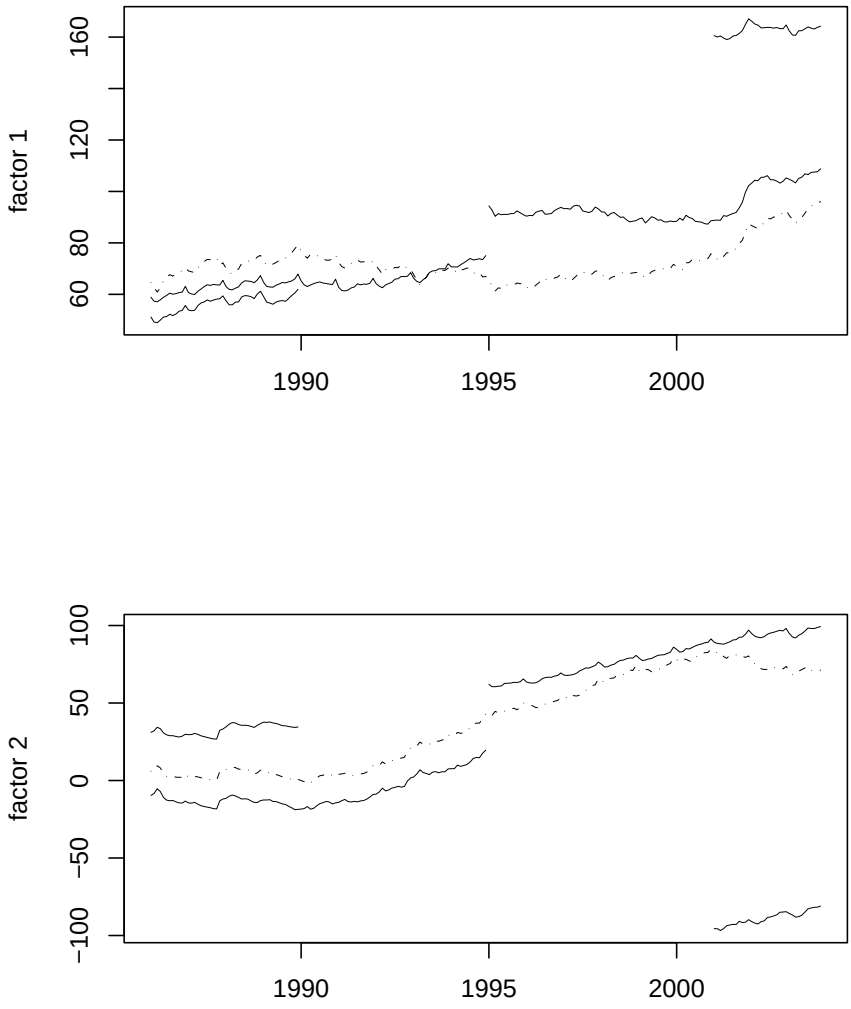

Figure 11: Bartlett predictors based on full sample and sub-samples.

due to variability in the loadings estimators. The differenced predictors are not shown (for space considerations) but are very similar on all subsamples, which suggests the technique is not overly sensitive to the selected sample. This is extremely encouraging, given that so much earlier work with the monetary aggregates over this period has suggested that structural change is an important issue.

7 Discussion

This paper presents a methodology which we have called *time series factor analysis* (TSFA) to distinguish it from *dynamic factor analysis* (DFA) and from standard (cross-sectional) factor analysis (FA). Compared to standard FA, data are differenced first, because time series data are often integrating. Moreover, a considerably weaker set of assumptions than for standard FA are shown to ensure parameter estimator consistency. Compared to DFA, the dynamics of the factors are not estimated. The reason for not doing this is to more clearly separate the measurement problem from the economic modeling problem.

Extracting the factors while making as few assumptions as possible about their dynamics can be important for a number of reasons. The most important is that specific assumptions about the dynamics of the model are usually much more fragile than the assumption that factors exist. When modeling and estimating both jointly, misspecification of the dynamic model for the factors may lead to distorted estimators for the measurement (factor analysis) part of the model. With TSFA, factors are established and measured fairly well before modeling their dynamics. TSFA may also be important in situations where one group (a statistics agency or central bank) measures the data and many researchers use it.

It is expected that TSFA is less efficient than DFA if the latter uses the correct specification of the dynamic process generating the factors. On the other hand, if DFA is based on a misspecified dynamic model, TSFA is expected to give better results. Hence, an important question is how large these losses and gains are. Some very preliminary simulations suggest that the efficiency loss of TSFA compared to a completely correctly specified DFA analysis may not be large, whereas DFA may give seriously biased results if the model for the factors is misspecified by omitting a seasonal component. A detailed study of the conditions under which TSFA is much better or worse than DFA is, however, beyond the scope of the current paper.

It is hoped that TSFA will measure factors that are better data for economic modeling than aggregates currently used. The TSFA and the subsequent economic modeling together would then be similar to DFA, but the TSFA step does not pre-suppose an economic model. Moreover, different economic models can be

compared more easily, because these use the same data.

The techniques developed for TSFA use standard factor analysis theory but are different in important respects: the data does not need to be covariance stationary (although the differenced data must satisfy a weak boundedness condition), there is a nontrivial mean structure, and the observations are allowed to be dependent over time. The effects of these on parameter estimation and factor score prediction have been discussed.

A striking empirical result is that the sampling variability in the parameters accounts for most of the errors in the factor score predictions, and the fundamental inability to estimate the factors consistently has only a very small impact for this type of data. However, the differenced factors appear to be estimated well, and in many respects this is even more important than estimating the levels precisely. If the indexes are pegged at a reasonable level then growth rates will be in a reasonable range.

Factor rotation still needs to be examined more closely. Results here are based on an oblimin rotation. This is appropriate if indicator data should load almost exclusively on one factor or the other. For several of the indicator series this may not be appropriate. Other rotation criteria will be considerd in future research.

The number of factors has been assumed here to be two, but there are indications that three factors may be needed. With only six indicators available, however, this gives a saturated, just-identified model that fits the observed covariance perfectly, regardless of the data. This leads to unreliable results and overfitting. This may best be accommodated by adding credit data. The Ledermann bound for six data series is three factors. Adding six credit series would give 12 in total, and a Ledermann bound between 7 and 8 factors. Since one would expect some factors to explain both credit and money behaviour, this seems like an attractive option.

There are also questions concerning structural breaks: whether they need to be considered, and if so, determining where they occur. Casual consideration of the results presented here suggests that structural breaks may not be a problem, but research on the monetary aggregates over this period has usually been frustrated by model breakdowns that are often attributed to structural breaks, so further investigation of this would be appropriate.

The estimation methodology should be applicable to a much larger range of problems than the application considered here, and indications are that it can work very well. The application to measuring money shows much promise and may provide a measure of a more fundamental underlying economic phenomena than conventional monetary aggregates.

References

T. W. Anderson. The use of factor analysis in the statistical analysis of multiple time series. *Psychometrika*, 28:1–25, 1963.

T. W. Anderson and Herman Rubin. Statistical inference in factor analysis. In J. Neyman, editor, *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability V*, pages 111–150. University of California Press, Berkeley, 1956.

Donald W. K. Andrews. Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica*, 59:817–858, 1991.

Claude O. Archer and Robert I. Jennrich. Standard errors for rotated factor loadings. *Psychometrika*, 38:581–592, 1973.

Jushan Bai. Inferential theory for factor models of large dimensions. *Econometrica*, 71: 135–171, 2003.

Jushan Bai and Serena Ng. Determining the number of factors in approximate factor models. *Econometrica*, 70:191–221, 2002.

Coen A. Bernaards and Robert I. Jennrich. Gradient projection algorithms and software for arbitrary rotation criteria in factor analysis. *Educational and Psychological Measurement*, in press.

T. Bollerslev. Generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics*, 31:307–327, 1986.

Michael W. Browne. An overview of analytic rotation in exploratory factor analysis. *Multivariate Behavioral Research*, 36:111–150, 2001.

R. B. Cattell. The description of personality I. Foundations of trait measurement. *Psychological Review*, 50:559–594, 1943.

R. B. Cattell, A. K. S. Cattell, and R. M. Rhymer. *P*-technique demonstrated in determining psycho-physiological source traits in a normal individual. *Psychometrika*, 12:267–288, 1947.

Gary Chamberlain and Michael Rothchild. Arbitrage, factor structure, and mean-variance analysis on large asset markets. *Econometrica*, 51:1281–1304, 1983.

Jakob De Haan, Erik Leertouwer, Erik Meijer, and Tom Wansbeek. Measuring central bank independence: A latent variables approach. *Scottish Journal of Political Economy*, 50:326–340, 2003.

Robert M. De Jong and James Davidson. Consistency of kernel estimators of heteroscedastic and autocorrelated covariance matrices. *Econometrica*, 68:407–423,

2000.

Mario Forni, Marc Hallin, Marco Lippi, and Lucrezia Reichlin. The generalized dynamic factor model: Identification and estimation. *The Review of Economics and Statistics*, 82:540–554, 2000.

John Geweke. The dynamic factor analysis of economic time-series models. In D. J. Aigner and A. S. Goldberger, editors, *Latent Variables in Socio-Economic Models*, pages 365–383. North-Holland, Amsterdam, 1977.

Paul D. Gilbert and Lise Pichette. Towards new money measures. *Money Affairs*, 15: 151–181, 2002.

Paul D. Gilbert and Lise Pichette. Dynamic factor analysis for measuring money. Working Paper 03-21, Bank of Canada, Ottawa, Canada, 2003. Also available at http://www.bank-banque-canada.ca/pgilbert/.

James D. Hamilton. State-space models. In Robert F. Engle and Daniel L. McFadden, editors, *Handbook of Econometrics*, volume IV, pages 3039–3080. Elsevier Science, Amsterdam, 1994.

Andrew C. Harvey. *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge University Press, Cambridge, UK, 1989.

Wayne H. Holtzman. Methodological issues in *P* technique. *Psychological Bulletin*, 59:248–256, 1962.

Robert I. Jennrich. Standard errors for obliquely rotated factor loadings. *Psychometrika*, 38:593–604, 1973.

Jeannie Kottaras. The construction of continuity-adjusted monetary aggregate components. Working Paper 22, Bank of Canada, Ottawa, Canada, 2003.

John C. Loehlin. *Latent Variable Models. An Introduction to Factor, Path, and Structural Analysis*. Erlbaum, Hillsdale, NJ, 1987.

Karl Löwner. Über monotone Matrixfunktionen. *Mathematische Zeitschrift*, 38: 177–216, 1934.

Erik Meijer and Tom Wansbeek. Quadratic prediction of factor scores. *Psychometrika*, 64:495–507, 1999.

Peter C. M. Molenaar. A dynamic factor model for the analysis of multivariate time series. *Psychometrika*, 50:181–202, 1985.

Whitney K. Newey and Kenneth D. West. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55:703–708, 1987.

R Development Core Team. *R: A language and environment for statistical*

computing. R Foundation for Statistical Computing, Vienna, Austria, 2004. URL `http://www.R-project.org`. ISBN 3-900051-07-0.

Alexander Shapiro. Asymptotic distribution theory in the analysis of covariance structures (a unified approach). *South African Statistical Journal*, 17:33–81, 1983.

Aris Spanos. Liquidity as a latent variable – An application of the MIMIC model. *Oxford Bulletin of Economics and Statistics*, 46:125–143, 1984.

James H. Stock and Mark W. Watson. Implications of dynamic factor models for VAR analysis. Working Paper 11467, National Bureau of Economic Research, Cambridge, MA, 2005.

Jos M. F. Ten Berge, Wim P. Krijnen, Tom Wansbeek, and Alexander Shapiro. Some new results on correlation preserving factor scores prediction methods. *Linear Algebra and its Applications*, 289:311–318, 1999.

Tom Wansbeek and Erik Meijer. *Measurement Error and Latent Variables in Econometrics*. North-Holland, Amsterdam, 2000.

Mark W. Watson and Robert F. Engle. Alternative algorithms for the estimation of dynamic factor, MIMIC and varying coefficient regression models. *Journal of Econometrics*, 23:385–400, 1983.

Halbert White. Maximum likelihood estimation of misspecified models. *Econometrica*, 50:1–25, 1982.