DE GRUYTER
OPEN

Journal of Official Statistics, Vol. 34, No. 1, 2018, pp. 265–301, http://dx.doi.org/10.1515/JOS-2018-0012

Factor Structural Time Series Models for Official Statistics with an Application to Hours Worked in Germany

Roland Weigand[1], *Susanne Wanger*[2], *and Ines Zapf*[3]

We introduce a high-dimensional structural time series model, where co-movement between the components is due to common factors. A two-step estimation strategy is presented, which is based on principal components in differences in a first step and state space methods in a second step. The methods add to the toolbox of official statisticians, constructing timely regular statistics from different data sources. In this context, we discuss typical measurement features such as survey errors, statistical breaks, different sampling frequencies and irregular observation patterns, and describe their statistical treatment. The methods are applied to the estimation of paid and unpaid overtime work as well as flows on working-time accounts in Germany, which enter the statistics on hours worked in the national accounts.

Key words: Unobserved components model; state space model; national accounts; overtime work; working-time accounts.

1. Introduction

In a very important and publicly visible field of official statistics, early releases of economic production or labor market indicators are constructed on a regular monthly or quarterly basis. Several surveys and other data sources are typically used to update these time series based on the information available so far. The importance of timely and precise measures of the economy is emphasized in a large literature on real-time data analysis, which shows that data revisions pose a severe challenge to forecasters and policymakers; see, for example, Croushore (2011). Hence, on the side of statistical agencies, most prominently for quarterly national accounts, efforts are made to produce accurate statistics by bringing together a large amount of primary data sources, typically surveys; see Bureau of Economic Analysis (2017), Wood and Elliott (2007), and Federal Statistical Office (2008) for GDP calculation in the US, in the UK, and in Germany.

The current article is a methodological contribution to this field of activity. For the estimation of a target series θ_t such as real GDP or hours worked in the past quarter, we make use not only of currently available surveys z_t that aim to measure θ_t, but notably also of the history of such surveys, $z_{t-1}, z_{t-2}, \ldots$, and of a possibly very large set of additional indicators, $x_t, x_{t-1}, \ldots$, which are in some way related to θ_t. Hence, in the terminology of survey or small area statistics, we discuss a new model and an

[1,2,3] Institute for Employment Research (IAB) Regensburger Strasse 104, 90478 Nuremberg, Germany. Emails: roland.weigand@posteo.de, susanne.wanger@iab.de, and ines.zapf@iab.de
Acknowledgments: The authors are grateful for helpful comments by Enzo Weber, by participants of the IAB-Bundesbank seminar 2013 in Frankfurt am Main and of the German Statistical Week 2014 in Hannover.

estimator that borrows strength both over time and from related variables in a data-rich environment.

Our proposed approach is based on factor models for high-dimensional time series which have become an indispensable tool for macroeconomic fore- and nowcasting as well as structural modeling; see Bai and Ng (2008) as well as Stock and Watson (2011) for recent surveys. In this field, seasonally adjusted variables typically enter the model in first or second differences, while the factors are modeled as a stationary VAR process. Methods for handling nonstationary variables are also available (Bai and Ng 2004), and unit-root versions of the factor-augmented VAR as well as error correction models are an area of active research; see, among others, Banerjee et al. (2014).

We propose a concurrent parametrization for large factor models of nonstationary variables which we formulate in the structural time series framework of Harvey (1991). Factor structures on the trend, seasonal, cyclical and irregular components allow to model the co-movements of a large number of time series in a parsimonious, componentwise manner. The popular common trends or common cycle models emerge as special cases, but a common features assumption, restricting the idiosyncratic part to be stationary or even serially uncorrelated, is not necessarily imposed in our framework. Rather, the idiosyncratic part may be characterized by trend, cycle and seasonal components as well.

For a straightforward and computationally feasible implementation of the approach, a principal component analysis is combined with state space methods in the spirit of Bräuning and Koopman (2014). We extract the principal components of suitably differenced data to account for nonstationarity of the idiosyncratic part. Re-cumulated factors are modeled jointly with the series of primary interest using likelihood-based techniques within a state space framework. In Monte Carlo simulations, we find that this method performs well, irrespectively of whether a common features assumption holds.

From the perspective of data construction, we discuss several advantages and possible modifications of our model in state space form. Primary sources in official statistics are typically subject to survey errors and statistical breaks. They may be collected at different sampling frequencies, while changing survey designs lead to irregular measurement patterns. Since the key part of our model is formulated in state space form, it is well-suited to handle these patterns. It produces efficient estimates of the target series when different surveys measure the same underlying series. Information from the past of the series is processed, and additional strength is borrowed from a large number of related series with correlated components. Seasonally adjusted time series, using all available data for the adjustment, are obtained as a by-product of the procedure.

The potential of the state space approach for official statistics has already been pointed out by other researchers. Uses in several areas of official statistics have been highlighted by Durbin (2000). There are examples where state space methods are applied for seasonal adjustment, while Pfeffermann (1991) and Tiller (1992) discuss signal extraction from repeated survey data. In small area statistics, state space models help obtain disaggregate figures from surveys by borrowing strength both over time and space, see Pfeffermann and Tiller (2006) and Krieg and van den Brakel (2012). In that context, Bollineni-Balabay et al. (2015) pursue the estimation of aggregates along with the small-area domains in the presence of survey redesigns and variance breaks. Durbin and Quenneville (1997) and Quenneville and Gagné (2013) introduce benchmark constraints drawn from precise but low-frequency census data to correct the preliminary survey estimates, while Harvey and Chung (2000) discuss modeling data from different sources, and Moauro and Savio (2005) is concerned with temporal disaggregation as required by national statistical agencies.

We apply our methodology to the statistics of hours worked in Germany. High-quality data on hours worked are a key for understanding aggregate labor market dynamics, for example, to track business cycles, to assess reactions to shocks such as the 2008/09 financial and economic crisis (Burda and Hunt 2011), and to confront macroeconomic theory with time series evidence (Ohanian and Raffo 2012). Timely figures on hourly labour productivity are considered as being important, for example, for well-guided wage negotiations and monetary policy.

In Germany, working time statistics are constructed within the working time measurement concept of the Institute for Employment Research (IAB). The componentwise accounts provide a comprehensive figure of hours worked and contributes results to the German national accounts; see Wanger et al. (2016). In the measurement of overtime hours and flows on working-time accounts (WTA), we use household and business surveys, while additionally drawing on several labor market and business cycle indicators. Lacking continuously available survey data on working-time account net flows, the latter is based on the unobserved trend and cycle components for transitory overtime hours as well as regular and actual hours worked.

The article is structured as follows: Section 2 describes the model and its statistical treatment, Section 3 illustrates alternative measurement schemes faced in official statistics and Section 4 presents finite sample properties of the procedure. Section 5 applies the methods to the German statistics of hours worked, while the last section concludes.

2. A High-Dimensional Structural Time Series Model

2.1. The Factor Model

This article presents a new model and its implementation for official statistics. It extends the scope of multivariate structural time series models (STSM) discussed by Harvey and Koopman (1997) to high-dimensional applications. As the point of departure, an N-dimensional vector time series y_t is decomposed into trend μ_t, seasonal γ_t, cycle c_t, and irregular components u_t, according to

$$y_t = \mu_t + \gamma_t + c_t + u_t, \tag{1}$$

where the terms on the right are unobserved stochastic processes. Additional components such as calendar effects or outliers can be straightforwardly incorporated through the use of dummy regressors given this additive formulation but are not considered in this article. After describing the dynamic specification of the components we introduce a factor structure that handles cross-series linkages within the groups of components, and the statistical treatment of the model.

We use a standard specification for the dynamics of each component and characterize the slow movements by local linear trends

$$\mu_{t+1} = \mu_t + \nu_t + \xi_t, \quad \nu_{t+1} = \nu_t + \zeta_t,$$

where $\xi_t \sim \text{iid } N(0, \Sigma_\xi)$ and $\zeta_t \sim \text{iid } N(0, \Sigma_\zeta)$ are independent Gaussian white noise sequences. For a model frequency of s observations per year, the seasonal components are

$$\gamma_{t+1} = -\sum_{j=0}^{s-2} \gamma_{t-j} + \omega_t, \quad \omega_t \sim \text{iid } N(0, \Sigma_\omega).$$

Alternatively, a trigonometric specification for the seasonal components can be used (Durbin and Koopman 2012, Sec. 3.2.1). An individual cycle component $\tilde{c}_{it}, i = 1, \ldots, N$ evolves jointly with the auxiliary process $\tilde{c}^*_{it}$ as

$$\begin{pmatrix} \tilde{c}_{i,t+1} \\ \tilde{c}^*_{i,t+1} \end{pmatrix} = \rho_i \begin{pmatrix} \cos\lambda_i & \sin\lambda_i \\ -\sin\lambda_i & \cos\lambda_i \end{pmatrix} \begin{pmatrix} \tilde{c}_{it} \\ \tilde{c}^*_{it} \end{pmatrix} + \begin{pmatrix} \kappa_{it} \\ \kappa^*_{it} \end{pmatrix}, \quad \begin{pmatrix} \kappa_{it} \\ \kappa^*_{it} \end{pmatrix} \sim \text{iid } N(0, \Sigma_{\kappa,ii} I),$$

where λ_i is the dominant frequency and $0 < \rho_i < 1$ denotes the dampening factor. As for the trends and seasonal components, linkages between the individual cycles are introduced through covariances between the disturbances, collected in Σ_κ. To gain flexibility on the temporal timing of the co-movement, we introduce phase shifts $\delta_2, \ldots, \delta_N$ between the cycles by setting $c_{it} = \tilde{c}_{it}\cos\lambda_i\delta_i + \tilde{c}^*_{it}\sin\lambda_i\delta_i$, $i = 1, \ldots, N$, where $\delta_1 = 0$ as a normalization and δ_i measures the lead time of cycle j against the cycle of the first variable; see Rünstler (2004) and Valle e Azevedo et al. (2006). Finally, the irregular noise term is given by $u_t \sim \text{iid } N(0, \Sigma_u)$. For simplicity we assume that all groups of shocks ξ_t, ζ_t, ω_t κ_t and u_t are mutually independent. Correlated components in the spirit of Morley et al. (2003) could be straightforwardly adapted as long as suitable identification conditions are met.

Our focus is on cases where N, the number of series in y_t is large, and hence a curse of dimensionality occurs in the unrestricted model (1). For full covariance matrices Σ_ξ, Σ_ζ, Σ_ω, Σ_κ and Σ_u, there are $O(N(N+1))$ variance parameters to be estimated, which makes the application practically infeasible even for moderate values of N. In such situations, factor models have been found useful for different purposes in economics and finance. They allow a parsimonious representation of the cross-section dependencies between panel units or time series variables. Within our STSM setup, we consider common factors for each group of components. Denoting the common components by a C superscript and the idiosyncratic terms by I, our model is given by

$$y_t = \Lambda_\mu \mu^C_t + \Lambda_\gamma \gamma^C_t + \Lambda_c c^C_t + \Lambda_u u^C_t + \mu^I_t + \gamma^I_t + c^I_t + u^I_t. \tag{2}$$

The common components are of dimensions r_μ, r_γ, r_c and r_u, respectively, which are typically substantially smaller than N, while Λ_k, $k \in \{\mu, \gamma, c, u\}$ are $N \times r_k$ loading matrices of full column rank. All components follow the same dynamics as those described below (1), and are driven by shocks with covariance matrices Σ^C_l and Σ^I_l for $l \in \{\xi, \zeta, \omega, \kappa, u\}$. The idiosyncratic components are assumed mutually uncorrelated and hence Σ^I_l are diagonal, so that the number of parameters is reduced to an order $O(N)$ for fixed factor dimensions.

Our decision of using a factor model to circumvent the curse of dimensionality is popular in the econometrics field, since this "reduced rank sparsity" can easily handle high correlations between the series due to business cycle linkages. Especially cyclical movements are candidates for such a rank reduction, but also long-term trends may be linked by identical underlying driving forces. The alternative way to avoid the dimensionality problem, the so-called "zero sparsity" where correlations are set to zero, is often less plausible in such setups. In the empirical application, we practice a mix of reduced rank and zero sparsity: Component variances and correlations are set to zero when this is statistically appropriate after the model dimension has been drastically reduced by the factor approach.

Identification of the latent components in (2) is achieved if two conditions hold: (a) The processes μ_t, γ_t, c_t and u_t from (1) are separately identified through their difference in dynamics, and (b) for each of these dynamic components, for example for $\Lambda_\mu \mu_t^C + \mu_t^I$, the common $\Lambda_\mu \mu_t^C$ and idiosyncratic μ_t^I are distinguished as in classical factor models by the assumption that the idiosyncratic series are mutually uncorrelated in the cross-section dimension. Condition (a) is standard both in univariate and multivariate structural time series models and unproblematic in the uncorrelated components case considered here. It is discussed among others by Harvey (1991, sc. 4.4). Condition (b) is not related to the dynamic properties of the series, but only draws on the correlations between the series which are due to a low-dimensional factor process. The autocorrelation and even nonstationarity, for example of $\Lambda_\mu \mu_t^C + \mu_t^I$, does not interfere with the identification problem since the setup can be easily transformed to the classical "white noise" factor model of Anderson (1984) by a univariate time series filter; the reversed filter applied to the identified components would in principle recover the original autocorrelation structure. Clearly, as in classical factor models, the loadings and factors are identified only up to rotation, so that the additional normalizations that the upper $r_k \times r_k$ block of suitable loading matrices are identity will be used in Subsection 2.2. The chosen identification, however, does not matter for the purpose of this article which is estimation of a target series rather than structural inference on the factors.

The factor STSM can be represented in the notation of a standard multivariate STSM (1) if a similar cycle assumption holds, that is, if all ρ_i and λ_i are identical for both the common and idiosyncratic components. However, the factor structure imposes restrictions on the disturbance covariance matrices, which are given by

$$\Sigma_\xi = \Lambda_\mu \Sigma_\xi^C \Lambda_\mu' + \Sigma_\xi^I, \quad \Sigma_\zeta = \Lambda_\mu \Sigma_\zeta^C \Lambda_\mu' + \Sigma_\zeta^I, \quad \Sigma_\omega = \Lambda_\gamma \Sigma_\omega^C \Lambda_\gamma' + \Sigma_\omega^I,$$
$$\Sigma_\kappa = \Lambda_c \Sigma_\kappa^C \Lambda_c' + \Sigma_\kappa^I, \quad \Sigma_u = \Lambda_u \Sigma_u^C \Lambda_u' + \Sigma_u^I.$$

If one or more of the columns of Λ_i are linearly dependent with those of Λ_j, $i \neq j$, the stacked loadings $(\Lambda_\mu, \Lambda_\gamma, \Lambda_c, \Lambda_u)$ have a reduced column rank denoted by $r < r_\mu + r_\gamma + r_c + r_u$. Then, the cross-section correlations between variables in y_t can be traced back to a smaller number of common sources than there are common structural time series components. This possibly smaller dimensional latent process is given by the r-dimensional compound factors denoted by f_t with a corresponding full-rank $N \times r$ loading matrix Λ, such that $y_t = \Lambda f_t + \mu_t^I + \gamma_t^I + c_t^I + u_t^I$. The compound factors are related to the common components by

$$f_t = \Gamma_\mu \mu_t^C + \Gamma_\gamma \gamma_t^C + \Gamma_c c_t^C + \Gamma_u u_t^C,$$

where $\Gamma_k = (\Lambda'\Lambda)^{-1}\Lambda'\Lambda_k$ are $r \times r_k$ matrices of full column rank. Again, factors f_t, loadings Λ and Γ_k are only identified up to linear combinations, but for a chosen rotation the common components μ_t^C, γ_t^C, c_t^C and u_t^C are identified (up to rotation) through their different dynamics, and hence can be estimated from f_t by the state space approach described by Harvey (1991). As an example with the richest dynamic structure possible for a given r, consider the case with $r = r_\mu = r_\gamma = r_c = r_u$ and $\Lambda = \Lambda_\mu = \Lambda_\gamma = \Lambda_c = \Lambda_u$. The factors $f_t = \mu_t^C + \gamma_t^C + c_t^C + u_t^C$ then follow a structural time series process and consist of trend, irregular, seasonal and cyclical components themselves.

Factor structures in the multivariate STSM framework have been studied before in the econometrics literature, albeit with a different scope. Models for a moderate number of series have been used to investigate common trends (and thus cointegration) or common cycles in their dynamics; see, for example, Harvey (1991, Sec. 8.5), Valle e Azevedo et al. (2006) or the software implementation of Koopman et al. (2009). In the standard setup, the idiosyncratic part is a white noise process, or at least has different dynamic properties from the factors'. Identification of the factor (e.g., μ_t^C in the common level model $y_t = \Lambda_\mu \mu_t^C + \varepsilon_t^I$) is therefore achieved in both of the ways (a) and (b) discussed above at the same time: In the common levels model μ_t^C is the only source of autocorrelation and also the only source of correlation between the series. This restricts the model in a very relevant way and makes it less applicable for high-dimensional settings, since especially in high dimensionsid iosyncratic errors with restricted dynamic properties (or even white noise) are unrealistic and a high factor dimension would be needed to provide a reasonable approximation to the data. Eickmeier (2009), among others, finds unit roots in the idiosyncratic part of many macroeconomic time series, so that a common trends assumption fails for a reasonable factor dimension.

Our Model (2), in contrast, allows factors and idiosyncratic part to have the same types of components as the common part, and hence to consist of trend, seasonal, cycle and noise. In this way, we may obtain a more parsimonious structure with less factors when a larger panel of data is considered. Our model is rather general in that it allows for co-movements in each of the components, while a common features restriction is possible by setting the respective idiosyncratic components, say trends or cycles, to zero. As we describe in the next section, our model allows a computationally feasible treatment even in the high-dimensional case, since it naturally allows a combination of PCA and state space methods. In contrast, in common cycles or common trends models the components are typically filtered out from a full state space approach which becomes cumbersome for a larger number of series and factors.

In the high-dimensional factor framework, by far the most popular approach for dynamic modeling is by estimating factors by principal components, and using VAR models for observed series and estimated factors, resulting in VAR-based dynamic factor or so-called factor-augmented VAR (FAVAR) models; see, for example Stock and Watson (2005) or Bernanke et al. (2005). Model (2) has several benefits also relative to such VAR-based approaches. Firstly, the structural approach offers insights into the nature of co-movements between the series, which can be assigned to specific components: Is it because of business cycles or rather correlated trends that macroeconomic time series co-move? Are there joint sources of changing seasonal patterns in several branches of the economy? Can common irregular components like weather effects be identified that transitorily hit several output measures? Secondly, in the context of filtering a signal from sparsely available data, the structural time series setup imposes a parsimonious parametrization which stabilizes the estimates. In the application to official statistics, all components help estimate the different features of the target series while taking into account all relevant information from related series. Thirdly, using information from many series may also lead to important improvements of seasonal adjustment procedures over univariate approaches.

2.2. *Estimation by Collapsing the Factor Space*

We suggest an estimation procedure of the model given by a combination of principal component and state space techniques along the lines of Bräuning and Koopman (2014). Assume that we are primarily interested in a low-dimensional subprocess z_t holding N_z series of the available data, while the complete set of time series is separated according to $y_t = (x_t', z_t')'$. In forecasting applications, z_t will hold at least the series to be predicted, while the estimation of official statistical figures typically requires the series z_t to consist of the major surveys measuring the target series. Unlike Bräuning and Koopman (2014), we assume that all variables in y_t are generated by the same model, (2) in our case, and hence variables in x_t and z_t are treated symmetrically in terms of the model but not in terms of the estimation procedure.

To estimate the space of compound factors f_t in a first step, we apply a suitable principal components analysis to x_t. By using the data x_t in differences, we avoid possible inconsistencies due to nonstationary idiosyncratic components, and thus adapt ideas of Bai and Ng (2004) to our setting. More concretely, denoting by L the lag operator, by $\Delta := (1 - L)$ the standard difference and by $\Delta_s := (1 - L_s)$ the seasonal difference operator, we obtain factor loadings $\bar{\Lambda}$ as $\sqrt{N_x}$ times the orthonormal eigenvectors corresponding to the r largest eigenvalues of $\sum_{t=1}^{T} (\Delta\Delta_s x_t)(\Delta\Delta_s x_t)'$. Estimated factors are obtained by re-cumulating the principal components in differences, or from the level data as $\bar{f}_t = \bar{\Lambda}' x_t$, which differs from the re-cumulation approach through the effects of initial values. Under an additional assumption on the factor loadings, the results of Bai and Ng (2002) are applicable to the variables in differences; see Appendix A. Among other things, this assures consistency (up to rotation and net of the effects of starting values) of $\bar{f}_t$ for f_t at a fixed t as N and T tend to infinity. In the setup (2), the differenced series are usually autocorrelated as are the residuals from the principal components approach. However, as long as differences of sufficient orders are taken, the autocorrelation is weak in the sense of Bai and Ng (2002, Assumption C) and consistency of the factors in this *approximate* factor framework is ensured.

The principal components approach is typically not optimal and comes with an efficiency loss, for example, in the situation of outliers due to nongaussianity, of heteroskedasticity of the idiosyncratic components, or of autocorrelation. We propagate its use as asimple, well-understood and popular first-step estimator, but of course improved versions are available and can also be applied in our setup (see, e.g., Breitung and Tenhofen 2011).

To gain information on the common components and their relation to the variables in z_t, we consider the joint model of f_t and z_t within the state space setup. Replacing the compound factors f_t by their estimates, the model is given by

$$\begin{pmatrix} \bar{f}_t \\ z_t \end{pmatrix} = \begin{pmatrix} \Gamma_\mu \\ \Lambda_\mu \end{pmatrix} \mu_t^C + \begin{pmatrix} \Gamma_\gamma \\ \Lambda_\gamma \end{pmatrix} \gamma_t^C + \begin{pmatrix} \Gamma_c \\ \Lambda_c \end{pmatrix} c_t^C + \begin{pmatrix} \Gamma_u \\ \Lambda_u \end{pmatrix} u_t^C + \begin{pmatrix} e_t \\ \mu_t^I + \gamma_t^I + c_t^I + u_t^I \end{pmatrix}, \quad (3)$$

where e_t is the error of $\bar{f}_t$ estimating f_t. As a slight abuse of notation, the idiosyncratic components and loadings are those corresponding to the elements in z_t only. While the compound factor estimates $\bar{f}_t$ are identified by the standard normalization of principal components, the common structural time series components are made unique by setting

$$\Gamma_\mu = \begin{pmatrix} I_{r_\mu} \\ \Gamma_\mu^{(2)} \end{pmatrix}, \quad \Gamma_\gamma = \begin{pmatrix} I_{r_\gamma} \\ \Gamma_\gamma^{(2)} \end{pmatrix}, \quad \Gamma_c = \begin{pmatrix} I_{r_c} \\ \Gamma_c^{(2)} \end{pmatrix}, \quad \Gamma_u = \begin{pmatrix} I_{r_u} \\ \Gamma_u^{(2)} \end{pmatrix},$$

and the common components may have unrestricted disturbance covariance matrices. Setting the upper block of the loading matrices to identity is only one of many ways to prohibit observationally equivalent rotations of factors and loadings (see e.g., Bai and Ng 2013), but especially fore- and nowcasts of the series do not depend on such normalizations.

Under the given restrictions, the model can be operationalized by ignoring the error from principal components estimation, and hence setting $e_t = 0$, which is justified as an approximation especially for large N. The unknown hyperparameters of (3) are estimated by maximum likelihood using the state space approach.

Alternatively, a multivariate STSM without the restrictions of (3) can be fitted to the joint process of principal components and variables of interest. This second strategy allows for correlation between the idiosyncratic terms of z_t, while the model nests the factor STSM specification (3). Given typical factor dimensions of less than five and a univariate or low-dimensional z_t, especially the latter estimation approach can easily be conducted in one of several available software packages such as STAMP (Koopman et al. 2009) or those described by Commandeur et al. (2011) and articles in the same special issue. For the computations in this article, the KFAS package for R is used (Helske 2016).

Empirically, the compound factor dimension can be inferred from the data y_t in suitable differences, for example by the methods proposed by Bai and Ng (2002). Alternatively, different (small) values of r can be considered and robustness with respect to this choice can be assessed in practice. Subsequently, for a given r, beginning from $r_\mu = r_\gamma = r_c = r_u = r$, the dimension of each common component may be determined in a general-to-specific sequential testing procedure based on (3).

3. Observation Schemes

The factor STSM introduced in this article has advantages in filtering latent series from incomplete measures which is a key issue in official statistics. For this purpose, we assume that a latent N_θ dimensional process θ_t of target series instead of observed z_t is modeled to follow the factor STSM (2), and that the observations collected in z_t are related to θ_t through a dynamic measurement relationship

$$z_t = d_t + M_t(L)\theta_t + \varepsilon_t, \quad \varepsilon_t \sim N(0, H_t), \quad (4)$$

where d_t holds possible survey bias terms and statistical breaks, $M_t(L) = M_{t0} + M_{t1}L + \ldots + M_{tl}L^l$ are $N_z \times N_\theta$ matrix lag polynomials holding the dynamic measurement coefficients, while ε_t is a vector of survey errors with possibly time-varying covariance matrices H_t. The latter need not necessarily follow a white noise process, but can, for example, contain autocorrelation due to survey overlap, which may be treated by the methods of Pfeffermann and Tiller (2006). We review some of the cases that the general mechanism (4) captures, and propose its implementation in the state space form which is given in Appendix B.

The measurement scheme (4) is sufficiently flexible to allow for several surveys estimating the same underlying concept, for missing data and for time-varying observation patterns. Consider an example where θ_{1t}, for example, paid overtime hours per week and employee, is measured by two surveys z_{1t} and z_{2t}, for example, the German Socio-Economic Panel (GSOEP), and the German Microcensus, as it is the case in the application to German hours worked data below. The measurement mechanism is then given by

$$z_{1t} = \theta_{1t} + \varepsilon_{1t}, \quad z_{2t} = d_2 + \theta_{1t} + \varepsilon_{2t}. \tag{5}$$

In this simple example, with $M_t(L) = (1, 1)'$, the scheme brings together contradicting surveys, where differences are explained by the survey errors ε_{1t} and ε_{2t}. The variances of these errors depend on the design and size of the survey and are likely to change over time. By including an unknown constant d_2 in the second measurement equations, it is possible to correct for a bias in one of the sources. Similarly, if statistical breaks, like changes in the survey questionnaire, occur in one or more of the data sources, these may be explicitly accounted for by level shifts in d_t, and hence leave the measured θ_t unaffected. In case of changing seasonal patterns or covariance structures of the components, however, a time-varying transition rather than measurement equation has to take this into account, a topic that we do not consider in this article.

Different sampling frequencies of regular surveys, or data missing for other reasons, are also covered by the measurement scheme (4), which is an important topic in the existing nowcasting literature (Giannone et al. 2008). Considering a quarterly stock variable which is measured only at the end of the quarter, we observe the monthly value $z_{1t} = \theta_{1t}$ only when t is the last month of a quarter, while values of z_t are missing two thirds of the time. Returning to the bivariate example, if in period t no survey z_{1t} is conducted, we obtain a trivial equation

$$0 = 0 \cdot \theta_{1t} + 0, \quad z_{2t} = d_2 + \theta_{1t} + \varepsilon_{2t} \tag{6}$$

by specifying $M_t(L) = (0, 1)'$ and $H_{11,t} := \mathrm{Var}(\varepsilon_{1t}) = 0$. Hence, no information is gained by the first survey in that period. Therefore, information about θ_{1t} stem firstly from other surveys z_{2t}, secondly from past and future values of z_{1t} through the dynamics of the system, or thirdly from additional indicators correlated with θ_{1t} through the common components. If one survey z_{1t} is used as a benchmark and hence the resulting estimate of θ_{1t} should exactly match that survey, this is reached by setting $\varepsilon_{1t} = 0$. Further relevant methods for benchmarking are discussed in Durbin and Quenneville (1997) and Quenneville and Gagné (2013).

In contrast to the previous example of a stock variable where survey interviews reflect observations on one period t, in reality reference intervals may span more than one period in terms of the model frequency. This is typically the case for flow variables like GDP where we observe the sum of the monthly flows $z_{1t} = \theta_t + \theta_{t-1} + \theta_{t-2}$ at the end of the quarter. With a lag polynomial $M(L) = 1 + L + L^2$, we can also write $z_{1t} = M(L)\theta_t$. As another example, since 2005 the German Microcensus has a continuous interview policy and allows an evaluation of quarterly averages of quantities such as overtime hours worked per week. If the model is formulated at monthly frequency, an observation z_{1t} of a flow variable, corresponding to the second quarter 2006, refers to the mean of the underlying θ_{1t}, $\theta_{1,t-1}$ and $\theta_{1,t-2}$ of April, May and June. The measurement equation reflects this by assigning the quarterly value z_{1t} to the last month of the quarter and selecting

$$z_{1t} = \frac{1}{3}\theta_{1t} + \frac{1}{3}\theta_{1,t-1} + \frac{1}{3}\theta_{1,t-2} = M_t(L)\theta_{1t} \tag{7}$$

where z_{1t} contains values only at the end of the quarter of each year, and where $M_t(L) = \frac{1}{3} + \frac{1}{3}L + \frac{1}{3}L^2$ is the lag polynomial that reflects that measurement scheme. The change from a fixed reference week to continuous interviews is reflected by a change in the time-varying observation polynomial $M_t(L)$, so that $M_t(L) = 1$ for periods t before 2005.

For other surveys, the observation scheme is still more general. For example, for household panel studies such as the GSOEP or the U.S. panel study of income dynamics (PSID), the field period spans several months and changes from year to year. Assigning the resulting yearly figure to the December of each survey year, an observation equation

$$z_{1t} = M_{t,dec}\theta_{1t} + \ldots + M_{t,jan}\theta_{1,t-11} \tag{8}$$

reflects the time-varying shares M_{tj} of observations in each month j, relative to all observations in that year. Figure 1 shows the distribution of the GSOEP interviews for

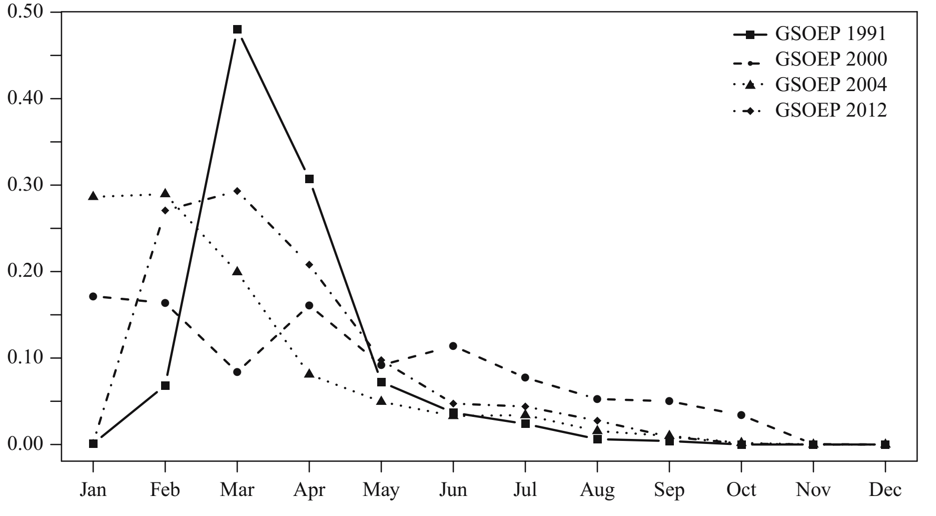

Fig. 1. Distribution of GSOEP interviews over certain years. The fraction of interviews for each month is shown for 1991, 2000, 2004, and 2012.

selected years, namely for 1991 (solid line), for 2000 (dashed), for 2004 (dotted), and for 2012 (dash-dotted).

Note that the measurement scheme might interfere with the identification of a given model. If an underlying series θ_{1t} is measured by a low-frequency, say quarterly, series alone, then seasonal components of some frequency may be unobservable; see Harvey (1991, Sec. 6.3). One obvious way to circumvent this is to use quarterly time-varying dummy variables that do not aim to estimate intra-quarter seasonality, while one could alternatively use a trigonometric seasonal specification (Durbin and Koopman 2012, Sec. 3.2.2) and skip frequencies higher than the observation frequency.

The Model (3) with the measurement scheme (4) can be stated in state space form (Appendix B). After estimating the model hyperparameters by the methods described in Subsection 2.2, estimated θ_t, $t = 1, \ldots, T$ using all available data are obtained by a state smoothing algorithm (Durbin and Koopman 2012, Sec. 4.3). The application of a smoother rather than a Kalman filter means that also past data are revised as soon as new information comes in. The smoother is constructed in a way that the revision optimally reflects the new information, given the model structure and its parameters. Since it is current practice in national accounts to revise also recent quarters, the use of a smoother automatically implements this revision together with the computation of a new quarter. Hence, no additional updating mechanism or model using more data is needed.

4. A Monte Carlo Study

A Monte Carlo study is conducted to shed light on the practical performance of the proposed methods in finite samples. Different aspects of the procedure are analysed. Firstly, the difference-based principal components approach is studied in the case of factor STSM processes for different data generating mechanisms and sample sizes, and compared to principal components in levels. Secondly, the estimation of the latent target process is evaluated and compared to standard benchmarks such as univariate models or standard principal-components based factor models.

4.1. Data Generating Processes

Four data generating processes are chosen to mimic different situations of practical relevance. We consider (1) cases with linearly independent loadings for the distinct common components and (2) cases with identical loadings, where principal components estimate a compound factor process. Furthermore, while typically (A) the idiosyncratic components have a structural time series structure with trend, seasonal and possibly cycle components, we additionally consider a common features assumption with (B) serially uncorrelated idiosyncratic components. We introduce the data generating processes as combinations of these characteristics in turn.

(1A) To define the first data generation mechanism as the case with linearly independent loadings and without common features, we consider the process (2) with $s = 4$ and where the cyclical components have frequency $\lambda = 0.2$ and dampening factor $\rho = 0.97$. The parameters in the loading matrices are randomly chosen for each draw. Denoting the uniform distribution between a and b by $U(a, b)$, they are given for $i = 1, \ldots, N$ and

$j = 1, \ldots, r$ by

$$\Lambda_{\mu,ij} \sim U(0,\chi), \quad \Lambda_{\gamma,ij} \sim U(0,\chi), \quad \Lambda_{c,ij} \sim U(0,\chi\varsigma), \quad \Lambda_{u,ij} \sim U(0,\chi\varsigma),$$

where the parameter $\mathcal{X}$ captures the overall importance of the common components relative to the idiosyncratic ones, while ς determines the relative size of stationary versus nonstationary components. The components are generated using innovation covariance matrices with

$$\Sigma^I_{\xi,ii} \sim U(0,1)^2, \quad \Sigma^I_{\zeta,ii} \sim U(0,1/10)^2, \quad \Sigma^I_{\omega,ii} \sim U(0,1)^2,$$

$$\Sigma^I_{\kappa,ii} \sim U(0,\varsigma)^2, \quad \Sigma^I_{u,ii} \sim U(0,\varsigma)^2$$

for idiosyncratic components and $\Sigma^C_\xi = 10\Sigma^C_\zeta = \Sigma^C_\omega = \Sigma^C_\kappa = \Sigma^C_u = I$ for common components, respectively.

(1B) The second data generating process is characterized by the same parameters for the common components as in (1A), but a common features assumption is imposed and hence the idiosyncratic components are subject to

$$\Sigma^I_\xi = \Sigma^I_\zeta = \Sigma^I_\omega = \Sigma^I_\kappa = 0, \quad \Sigma^I_{u,ii} \sim U(0,5)^2.$$

(2A) To introduce cases with linearly dependent common components loadings, the third data generating process sets the compound loading matrix Λ according to

$$\Lambda_{\mu,ij} = \Lambda_{\gamma,ij} = \frac{1}{\varsigma}\Lambda_{u,ij} \sim U(0,\chi),$$

and drops the cyclical components from the processes. The remaining variances Σ^I_ξ, Σ^I_ζ, Σ^I_ω, Σ^I_u, Σ^C_ξ, Σ^C_ζ, Σ^C_ω, and Σ^C_u correspond to those in (1A).

(2B) The last data generating process drops the trend and seasonal from the idiosyncratic components of the previous one, so that the only difference to (2A) is in the covariance matrices

$$\Sigma^I_\xi = \Sigma^I_\zeta = \Sigma^I_\omega = 0, \quad \Sigma^I_{u,ii} \sim U(0,5)^2.$$

4.2. *Estimation of the Compound Factor Space*

We first assess the performance of the principal components procedure based on differenced data $\Delta\Delta_4 y_t$ which we have proposed as a first step in estimating the factor STSM. For all data generating processes and different values for the time and cross-section dimensions T and N, we simulate 1,000 trajectories and repeatedly estimate the compound factor process f_t by $\bar{f}_t$ as explained in Subsection 2.2. We compare the results to the principal component method using the data in levels. At this step, no maximum likelihood estimation of the structural model is performed and hence the dynamic properties are not taken into account. Therefore, only the space of compound factors can be estimated, which is identified only up to rotation. The measure of estimation error has to take this lack of identification into account, and hence we rotate each factor estimate to achieve the best predictive power for the true factors by least squares. Since the overall level of mean squared error is very different across estimated components and has no natural interpretation, we standardize the mean squared error by reporting the adjusted R^2 from regressing each of the true compound factors on the estimated factors $\bar{f}_t$. This measure is strictly decreasing in the (root) mean squared error and emphasizes the size of the error relative to the overall factor variation. To enforce stationarity for these evaluation equations, we apply the regressions in differences ($\Delta_4\Delta$) of the true factors and their estimates. These R^2 are averaged over the iterations.

Table 1 gives results for the data generating process (1A) with linearly independent component loadings and without a common components structure. There, the true factor process consists of the $r = 4$ common structural times series components, $f_t = (\mu_t^C, \gamma_t^C, c_t^C, u_t^C)'$, which allows for an evaluation of each component separately. Overall, the principal components in differences outperform the estimates based on levels data. The difference between the methods is most pronounced for larger N and T. The estimates in differences clearly improve with N, but also slightly with T, with R^2 becoming close to one for each component in large samples. The level estimate, however, especially the stationary components in the baseline case with $\chi = 1$ and $\varsigma = 1$, does not show a clear improvement with larger N. The precision typically even worsens with larger T, which is the result of inconsistency when the idiosyncratic components are nonstationary; see Bai and Ng (2004) for the I(1) framework. The results are robust to changing the scale of the stationary components to $\varsigma = 2$ and of the common factors to $\chi = 2$. These changes lead to the expected results that the stationary common components are estimated more precisely in the former case, while the overall precision increases in the latter case.

In Table 2, we show results for the process (1B) which entails the common features assumption that the idiosyncratic components are white noise. Compared to (1A), the overall picture changes. Now, the estimates in levels are better than their difference-based counterparts, most strikingly for larger N. The difference-based estimates still improve both with N and with T. The precisions of the two estimators for the stationary components are closer to each other for $\varsigma = 2$ and for $\chi = 2$, but still the level-based estimates dominate the difference-based ones almost uniformly.

Results for the data generating processes with identical loadings for all common components are depicted in Table 3. We evaluate the precision of r principal components estimating the r compound factors $f_t = \mu_t^C + \gamma_t^C + u_t^C$ for $r \in \{1, 2\}$ by means of the adjusted R^2 as before. The mean of the adjusted R^2 over both evaluation regressions is computed in the case $r = 2$.

For $r = 1$, the adjusted R^2 are very close to one for all chosen specifications. Thus, when compared to Tables 1 and 2, the performance is seemingly enhanced if the components can be estimated in aggregated form, which reduces the compound factor dimension relative to the first two data generating processes. However, the higher uncertainty of the first two cases likely recurs in the second step when distinct structural time series components are estimated from the compound factors in a state space framework. The outcomes for $r = 2$ reveal a loss of precision and visible differences between the specifications and estimators. The patterns described for the first two data generating processes are confirmed here. Most notably, without the common feature assumption the difference estimator outperforms the level estimator again, while the latter is slightly better in case of common features.

Table 1. Precision of common component estimation by principal components in levels and differences ($\Delta_4\Delta$) for process (1A) without common features. The mean of the adjusted R^2 from regressions of true common components on estimated factors is given.

				pca in levels				pca in differences			
χ	ς	T	N	μ_t	γ_t	c_t	u_t	μ_t	γ_t	c_t	u_t
1	1	250	10	0.260	0.506	0.270	0.338	0.229	0.551	0.310	0.430
1	1	250	50	0.427	0.633	0.272	0.338	0.507	0.836	0.637	0.756
1	1	250	100	0.499	0.688	0.270	0.333	0.736	0.917	0.812	0.873
1	1	250	500	0.634	0.780	0.278	0.338	0.941	0.983	0.959	0.973
1	1	500	10	0.260	0.496	0.253	0.333	0.232	0.548	0.306	0.418
1	1	500	50	0.369	0.588	0.260	0.336	0.564	0.846	0.676	0.773
1	1	500	100	0.427	0.638	0.260	0.346	0.770	0.922	0.830	0.884
1	1	500	500	0.550	0.720	0.257	0.338	0.951	0.984	0.964	0.976
1	1	1000	10	0.254	0.469	0.257	0.339	0.230	0.548	0.317	0.422
1	1	1000	50	0.309	0.504	0.262	0.348	0.588	0.853	0.698	0.785
1	1	1000	100	0.316	0.503	0.260	0.352	0.783	0.924	0.839	0.888
1	1	1000	500	0.347	0.520	0.265	0.355	0.953	0.985	0.966	0.977
1	2	250	10	0.136	0.287	0.445	0.499	0.096	0.291	0.483	0.603
1	2	250	50	0.254	0.366	0.550	0.516	0.147	0.627	0.799	0.861
1	2	250	100	0.336	0.414	0.584	0.522	0.260	0.783	0.889	0.924
1	2	250	500	0.517	0.516	0.655	0.535	0.806	0.951	0.976	0.984
1	2	500	10	0.127	0.277	0.420	0.499	0.094	0.288	0.490	0.603
1	2	500	50	0.209	0.343	0.439	0.523	0.157	0.655	0.808	0.867
1	2	500	100	0.259	0.371	0.440	0.525	0.369	0.807	0.898	0.928
1	2	500	500	0.425	0.459	0.448	0.532	0.864	0.959	0.979	0.985
1	2	1000	10	0.125	0.263	0.403	0.502	0.091	0.289	0.483	0.600
1	2	1000	50	0.151	0.275	0.407	0.528	0.167	0.668	0.813	0.869
1	2	1000	100	0.160	0.274	0.420	0.528	0.453	0.819	0.901	0.931
1	2	1000	500	0.180	0.294	0.418	0.535	0.883	0.962	0.980	0.986
2	1	250	10	0.492	0.716	0.487	0.473	0.497	0.783	0.605	0.695
2	1	250	50	0.717	0.841	0.514	0.453	0.878	0.959	0.911	0.939
2	1	250	100	0.767	0.869	0.537	0.447	0.941	0.980	0.956	0.970
2	1	250	500	0.825	0.901	0.628	0.422	0.988	0.996	0.991	0.994
2	1	500	10	0.482	0.717	0.417	0.484	0.500	0.788	0.603	0.695
2	1	500	50	0.668	0.820	0.387	0.465	0.884	0.959	0.914	0.940
2	1	500	100	0.714	0.845	0.373	0.463	0.942	0.980	0.957	0.970
2	1	500	500	0.764	0.872	0.378	0.460	0.988	0.996	0.992	0.994
2	1	1000	10	0.453	0.682	0.413	0.493	0.514	0.791	0.609	0.700
2	1	1000	50	0.546	0.726	0.387	0.480	0.885	0.960	0.916	0.940
2	1	1000	100	0.566	0.740	0.396	0.480	0.943	0.980	0.958	0.971
2	1	1000	500	0.620	0.773	0.378	0.466	0.989	0.996	0.992	0.994

These outcomes suggest that the estimator choice should be based on whether the idiosyncratic components are white noise or not, and that unnecessary differencing should be avoided. The difference-based estimator appears as a robust choice since it is consistent in both settings while the level-based estimator does not necessarily improve with sample size in the general framework of this article.

Table 2. Precision of common component estimation by principal components in levels and differences ($\Delta_4\Delta$) for process (1B) with common features. The mean of the adjusted R^2 from regressions of true common components on estimated factors is given.

				pca in levels				pca in differences			
χ	ς	T	N	μ_t	γ_t	c_t	u_t	μ_t	γ_t	c_t	u_t
1	1	250	10	0.150	0.331	0.160	0.185	0.079	0.230	0.101	0.163
1	1	250	50	0.300	0.556	0.343	0.237	0.120	0.381	0.157	0.263
1	1	250	100	0.414	0.673	0.468	0.275	0.136	0.440	0.177	0.305
1	1	250	500	0.756	0.900	0.802	0.742	0.176	0.631	0.234	0.519
1	1	500	10	0.145	0.331	0.154	0.180	0.075	0.221	0.098	0.156
1	1	500	50	0.297	0.559	0.343	0.234	0.117	0.383	0.156	0.252
1	1	500	100	0.410	0.677	0.471	0.276	0.132	0.444	0.179	0.298
1	1	500	500	0.763	0.905	0.810	0.794	0.199	0.741	0.302	0.618
1	1	1000	10	0.143	0.338	0.158	0.175	0.073	0.220	0.100	0.148
1	1	1000	50	0.296	0.564	0.342	0.231	0.115	0.382	0.159	0.246
1	1	1000	100	0.414	0.678	0.475	0.287	0.133	0.448	0.183	0.300
1	1	1000	500	0.767	0.907	0.816	0.818	0.247	0.821	0.460	0.710
1	2	250	10	0.110	0.249	0.386	0.402	0.057	0.163	0.281	0.411
1	2	250	50	0.257	0.494	0.651	0.606	0.069	0.213	0.412	0.566
1	2	250	100	0.403	0.659	0.785	0.785	0.074	0.272	0.535	0.688
1	2	250	500	0.767	0.905	0.947	0.957	0.101	0.653	0.843	0.918
1	2	500	10	0.109	0.252	0.391	0.399	0.054	0.158	0.285	0.403
1	2	500	50	0.261	0.506	0.654	0.630	0.063	0.211	0.426	0.577
1	2	500	100	0.402	0.666	0.790	0.801	0.069	0.288	0.595	0.732
1	2	500	500	0.768	0.906	0.947	0.959	0.112	0.786	0.901	0.940
1	2	1000	10	0.109	0.250	0.388	0.394	0.051	0.148	0.279	0.396
1	2	1000	50	0.260	0.509	0.658	0.645	0.060	0.208	0.440	0.589
1	2	1000	100	0.404	0.668	0.791	0.808	0.068	0.330	0.639	0.755
1	2	1000	500	0.768	0.908	0.948	0.960	0.194	0.835	0.922	0.949
2	1	250	10	0.286	0.541	0.331	0.299	0.162	0.442	0.216	0.315
2	1	250	50	0.565	0.792	0.633	0.534	0.213	0.636	0.303	0.499
2	1	250	100	0.716	0.884	0.774	0.764	0.288	0.784	0.437	0.663
2	1	250	500	0.929	0.974	0.946	0.955	0.759	0.953	0.850	0.926
2	1	500	10	0.281	0.540	0.337	0.296	0.156	0.433	0.216	0.306
2	1	500	50	0.571	0.793	0.632	0.564	0.217	0.656	0.322	0.512
2	1	500	100	0.721	0.885	0.776	0.783	0.329	0.818	0.519	0.710
2	1	500	500	0.929	0.975	0.946	0.958	0.868	0.966	0.911	0.947
2	1	1000	10	0.282	0.543	0.332	0.291	0.151	0.426	0.213	0.304
2	1	1000	50	0.571	0.799	0.635	0.582	0.225	0.677	0.347	0.527
2	1	1000	100	0.725	0.887	0.779	0.791	0.407	0.838	0.593	0.743
2	1	1000	500	0.930	0.975	0.947	0.959	0.900	0.970	0.931	0.954

4.3. Estimation of Latent Processes

In a second part of this Monte Carlo study, we assess the two-step procedure with respect to its ability to estimate a latent process θ_{1t} from incomplete data z_{1t} and additional high-dimensional data x_t. We delete $N/4$ of the observations in z_{1t} which is generated together

Table 3. Precision of compound factor estimation by principal components in levels and differences ($\Delta_4\Delta$) for processes (2A) and (2B) (without and with common features). The mean of the adjusted R^2 from regressions of true common components on estimated factors is given.

				$r=1$				$r=2$			
				(2A)		(2B)		(2A)		(2B)	
χ	ς	T	N	level	diff	level	diff	level	diff	level	diff
1	1	250	10	0.970	0.994	0.998	0.985	0.644	0.760	0.463	0.283
1	1	250	50	0.973	0.999	1.000	0.999	0.798	0.951	0.752	0.494
1	1	250	100	0.969	0.999	1.000	1.000	0.843	0.976	0.850	0.614
1	1	250	500	0.978	1.000	1.000	1.000	0.932	0.995	0.964	0.919
1	1	500	10	0.977	0.995	0.998	0.986	0.628	0.764	0.469	0.279
1	1	500	50	0.983	0.999	1.000	0.999	0.772	0.952	0.753	0.507
1	1	500	100	0.987	0.999	1.000	1.000	0.825	0.976	0.850	0.684
1	1	500	500	0.990	1.000	1.000	1.000	0.915	0.995	0.964	0.945
1	1	1000	10	0.976	0.994	0.998	0.985	0.616	0.767	0.465	0.280
1	1	1000	50	0.990	0.999	1.000	1.000	0.761	0.953	0.754	0.527
1	1	1000	100	0.992	0.999	1.000	1.000	0.809	0.977	0.850	0.733
1	1	1000	500	0.996	1.000	1.000	1.000	0.907	0.995	0.964	0.954
1	2	250	10	0.965	0.994	0.998	0.992	0.642	0.746	0.586	0.414
1	2	250	50	0.972	0.999	1.000	1.000	0.801	0.951	0.850	0.694
1	2	250	100	0.978	0.999	1.000	1.000	0.843	0.975	0.916	0.857
1	2	250	500	0.974	1.000	1.000	1.000	0.928	0.995	0.981	0.972
1	2	500	10	0.970	0.993	0.998	0.995	0.627	0.750	0.593	0.414
1	2	500	50	0.985	0.999	1.000	1.000	0.772	0.952	0.847	0.731
1	2	500	100	0.987	0.999	1.000	1.000	0.829	0.976	0.916	0.878
1	2	500	500	0.991	1.000	1.000	1.000	0.913	0.995	0.981	0.977
1	2	1000	10	0.976	0.993	0.998	0.995	0.615	0.756	0.593	0.413
1	2	1000	50	0.986	0.999	1.000	1.000	0.759	0.952	0.849	0.756
1	2	1000	100	0.991	0.999	1.000	1.000	0.807	0.976	0.916	0.888
1	2	1000	500	0.996	1.000	1.000	1.000	0.905	0.995	0.981	0.979
2	1	250	10	0.985	0.999	0.999	0.999	0.867	0.930	0.710	0.538
2	1	250	50	0.990	1.000	1.000	1.000	0.948	0.988	0.914	0.872
2	1	250	100	0.988	1.000	1.000	1.000	0.965	0.994	0.955	0.941
2	1	250	500	0.988	1.000	1.000	1.000	0.985	0.999	0.991	0.988
2	1	500	10	0.991	0.998	1.000	0.999	0.844	0.931	0.708	0.540
2	1	500	50	0.994	1.000	1.000	1.000	0.933	0.988	0.915	0.885
2	1	500	100	0.997	1.000	1.000	1.000	0.955	0.994	0.955	0.945
2	1	500	500	0.998	1.000	1.000	1.000	0.982	0.999	0.991	0.990
2	1	1000	10	0.993	0.998	1.000	0.999	0.837	0.929	0.711	0.544
2	1	1000	50	0.996	1.000	1.000	1.000	0.928	0.988	0.915	0.890
2	1	1000	100	0.998	1.000	1.000	1.000	0.949	0.994	0.955	0.948
2	1	1000	500	0.998	1.000	1.000	1.000	0.980	0.999	0.991	0.990

with x_t as a factor STSM for $y_t' = (z_{1t}, x_t')$. Different alternative approaches are considered to estimate θ_{1t} for each observation where z_{1t} is missing. As an infeasible benchmark, (i) the factor STSM with known factor process f_t is considered in state space form, where parameters are determined by maximum likelihood and missing values are estimated by the state smoother. As the feasible counterpart, (ii) the two-step estimate proposed in this article is used, where f_t is estimated by the principal components based on data in differences $\Delta\Delta_4 x_t$. As one further straightforward benchmark we use (iii) a univariate STSM which neglects information from x_t. Comparison between (ii) and (iii) straightforwardly illustrates the effect of taking into account the common part θ_t^C versus ignoring it.

As a simple competitor that also uses time series information on z_{1t} only, we interpolate the series using (iv) a local mean of available $\Delta_4 z_{1t}$ in the range of ± 20 observations near the period to be estimated. Cross-section information, but not the dynamics of the system are utilized by static regression-based predictions of θ_{1t} using the difference-based principal components of x_t as predictors. The regression is run (v) in levels, (vi) applying a yearly difference operator Δ_4 to z_{1t} and the principal component, or (vii) applying the difference operator $\Delta\Delta_4$ which is sufficient to make the variables stationary. A comparison to the full state space model, possibly using the common features restriction, would allow a measure of the undergone efficiency loss by our method but is beyond the scope of this article: The high dimension makes the treatment computationally intractable both here and in similar empirical problems, so that we omit it from this comparison.

Table 4 shows the corresponding root mean squared errors (RMSE) from estimating θ_{1t} according to the data generating process (1B) with $r = 1$. Not surprisingly, the infeasible estimator (i) outperforms the others, while the feasible two-step strategy (ii) of utilizing the factor STSM comes a close second. The loss from having to estimate f_t is rather small in this specification, and amounts to less than five percent of the overall RMSE in most cases. Clearly, this result may strongly depend on the data generating process and the corresponding precision of the principal components method. The differences vanish with larger N.

The univariate STSM approach (iii) comes in third place, but missing information on the factors leads to an efficiency loss which is more pronounced if either $\varsigma = 2$ which increases the noise which is unpredictable by univariate methods, or if $\chi = 2$ where the information content of x_t is higher. However, taking the dynamics into account appropriately pays off, which turns out from a comparison to the naïve local averaging method that performs clearly worse than all STSM approaches. The static regression estimation with principal components as predictors (v) leads to very spurious results in levels, while it still does not lead to a relevant improvement even over the local averaging method when it is applied in differences (vi, vii).

5. Application to German Hours Worked Statistics

We apply the proposed techniques to the measurement of several components of hours worked in Germany. Official statistics on hours worked per person and the overall volume of work are determined by the IAB which contributes the corresponding time series to the German national accounts. The working time measurement concept is a componentwise system where collective, calendar, cyclical, personal and other components are determined separately on a quarterly basis since 1991, and results are disaggregated according to industries, regions, and employment status; see Wanger (2013) and Wanger et al. (2016) for recent overviews.

Table 4. RMSE of estimating N/4 randomly chosen missing values for z_{1t} in process (2A) with $r = 1$. (i) Factor STSM with known factors f_t, (ii) factor STSM with differenced pca factor estimates, (iii) univariate STSM, (iv) mean of yearly differences of z_{1t} within ± 20 observations, (v)–(vii) static OLS of z_{1t} on pca in levels and differences.

χ	ς	T	N	STSM f_t	STSM pca	STSM univar.	Mean Δ_4	OLS level	OLS Δ_4	OLS $\Delta\Delta_4$
1	1	250	10	1.126	1.210	1.627	2.518	4.743	2.217	2.490
1	1	250	50	1.148	1.166	1.653	2.550	4.772	2.189	2.411
1	1	250	100	1.151	1.159	1.634	2.512	4.584	2.133	2.416
1	1	250	500	1.160	1.161	1.663	2.558	4.575	2.136	2.397
1	1	500	10	1.187	1.261	1.644	2.548	10.621	2.642	2.555
1	1	500	50	1.181	1.193	1.612	2.502	9.964	2.498	2.420
1	1	500	100	1.180	1.187	1.644	2.567	9.733	2.498	2.393
1	1	500	500	1.165	1.166	1.598	2.501	9.803	2.454	2.359
1	1	1000	10	1.275	1.328	1.636	2.567	26.310	3.195	2.562
1	1	1000	50	1.255	1.265	1.614	2.526	24.397	2.955	2.388
1	1	1000	100	1.268	1.274	1.630	2.554	24.386	2.986	2.386
1	1	1000	500	1.267	1.270	1.624	2.549	24.330	2.970	2.368
1	2	250	10	1.522	1.650	2.245	3.182	4.899	2.649	3.310
1	2	250	50	1.556	1.582	2.283	3.224	4.923	2.597	3.203
1	2	250	100	1.565	1.576	2.254	3.187	4.728	2.561	3.209
1	2	250	500	1.568	1.570	2.290	3.238	4.727	2.555	3.194
1	2	500	10	1.569	1.692	2.274	3.238	10.705	3.064	3.428
1	2	500	50	1.546	1.568	2.220	3.158	10.062	2.881	3.198
1	2	500	100	1.545	1.557	2.272	3.243	9.826	2.878	3.166
1	2	500	500	1.543	1.545	2.212	3.167	9.889	2.836	3.139
1	2	1000	10	1.595	1.700	2.259	3.246	26.281	3.561	3.400
1	2	1000	50	1.555	1.575	2.220	3.183	24.436	3.292	3.149
1	2	1000	100	1.573	1.583	2.241	3.218	24.424	3.328	3.151
1	2	1000	500	1.587	1.589	2.252	3.232	24.356	3.323	3.154
2	1	250	10	1.200	1.277	2.514	3.837	4.756	2.225	2.496
2	1	250	50	1.232	1.248	2.543	3.886	4.768	2.189	2.412
2	1	250	100	1.221	1.229	2.497	3.800	4.577	2.131	2.414
2	1	250	500	1.244	1.247	2.563	3.909	4.575	2.136	2.397
2	1	500	10	1.404	1.466	2.534	3.884	10.652	2.654	2.563
2	1	500	50	1.384	1.394	2.456	3.781	9.972	2.502	2.422
2	1	500	100	1.395	1.402	2.542	3.915	9.732	2.498	2.393
2	1	500	500	1.357	1.361	2.440	3.771	9.811	2.455	2.358
2	1	1000	10	1.683	1.719	2.512	3.893	26.520	3.209	2.571
2	1	1000	50	1.664	1.673	2.488	3.849	24.408	2.955	2.388
2	1	1000	100	1.677	1.683	2.518	3.901	24.416	2.985	2.385
2	1	1000	500	1.687	1.692	2.505	3.888	24.327	2.969	2.368

During a major revision in 2014 which also affected the methodology of the working time measurement, state space techniques were introduced to enhance the estimation precision for components with incomplete data sources, or where more than one source is used in the measurement. In this section, we describe the computation of the cyclical components paid and unpaid overtime work as well as flows on working-time accounts. These are of primary importance when assessing the business cycle fluctuations of hours worked in real time.

5.1. Paid and Unpaid Overtime Hours

The computations of overtime hours in the working-time measurement concept are primarily based on two yearly surveys. In the GSOEP, employed persons are asked for the number of performed overtime hours in the recent month and the way overtime work is typically compensated. From the responses, yearly time series on paid and unpaid overtime hours since the 1980s can be constructed, but as has been mentioned in Section 3, a changing distribution of interviews over the year has to be taken into account when considering a target series of higher frequency. As a second primary data source, the Microcensus offers information on paid and unpaid overtime hours since 2010 on the basis of quarterly averages.

The main problem of constructing a quarterly time series in real time is the substantial publication lag of each of the sources, since results from the GSOEP are available approximately twelve months after the end of a reference year, while the Microcensus results typically come in July of the following year. Hence, information regarding the first quarter of each year is available only after about 21 months (GSOEP) and 16 months (Microcensus), respectively. Additionally, the determination of intra-year fluctuations before 2010 is challenging, since until then only yearly GSOEP data are available. In response, we gather additional indicators to tackle these problems and to achieve the highest possible precision for the given available data.

As an additional data source, we consider the Ifo Business Survey from the Ifo Institute (Leibniz Institute for Economic Research at the University of Munich). In this survey, establishments are asked in the last month of each quarter whether their employees currently perform overtime work. Along with the log of the GSOEP and of the Microcensus measures of overtime hours per week (z_{1t} and z_{2t}, respectively), the logarithmic fraction of establishments with overtime work enters the model as a third series of interest, z_{3t}.

Further economic and labor market indicators (x_t) are used to compute principal components which enter the factor STSM. Here, we use real gross domestic product, the production index, new orders for all manufacturing industries, the number of employed persons, real compensation per employee (all from the Federal Statistical Office), registered unemployment (from the Federal Employment Agency), business expectations, business assessment and the employment barometer (from the Ifo Institute) as well as the willingness to buy index (from GfK Nuremberg). These variables are considered informative when assessing the current business cycle and labor market development, and hence for the amount of overtime work. We refrain from using a data set of higher dimension, since the additional data are likely to introduce irrelevant information and require a higher number of factors. At the same time, we keep the updating process simple by this choice.

Principal component estimates are computed after applying the natural logarithm to all variables except business expectations, business assessment, the employment barometer and the willingness to buy index. Seasonally adjusted data are used in x_t, so that yearly differences are not needed to remove seasonal nonstationarity. Additionally, there is no evidence for a changing slope in the processes: The p-values for tests of $\Sigma_\zeta = 0$ in univariate STSM is 0.97 and 0.07 for the first and second principal component (based on second differences), respectively. Hence, we base the subsequent analysis on re-cumulated principal components from first differences of the raw data. Data gaps and mixed frequency issues in x_t are resolved by the algorithm described by Stock and Watson (2002).

The resulting models for paid and unpaid overtime hours, respectively, are formulated in terms of a monthly model frequency to precisely capture the timing of the measurement process. Along with the r estimated factors $\bar{f}_t$, which capture the compound common components θ_t^C on a monthly basis, the measurement model is given by

$$\begin{pmatrix} \bar{f}_t \\ \log(ot\_gsoep_t) \\ \log(ot\_mc_t) \\ \log(ot\_ifo_t) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ d_2 \\ 0 \end{pmatrix} + \begin{pmatrix} I & 0 & 0 \\ 0 & M_{11,t}(L) & 0 \\ 0 & M_{21,t}(L) & 0 \\ 0 & 0 & M_{33,t}(L) \end{pmatrix} \begin{pmatrix} \theta_t^C \\ \theta_{1t} \\ \theta_{2t} \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \varepsilon_{2t} \\ 0 \end{pmatrix}.$$

The model comprises a monthly variable $\bar{f}_t$, a yearly ot_gsoep_t which is brought into the model at the last month of the year only, the quarterly ot_mc_t, which is brought in at the last month of the quarter, and the likewise quarterly ot_ifo_t, which refers to a single month of each quarter where it comes into the model. The GSOEP measurement scheme $M_{11,t}(L)$ is determined by the changing proportion of GSOEP-interviews in each month as in (8) so that $M_{11,t}(L) = M_{11,t,dec} + M_{11,t,nov}L + \ldots + M_{11,t,jan}L^{11}$. As an example for the year 2013, the obsevation of ot_gsoep_t is given for $t = 2013M12$, while all other months are missing. The proportion of interviews accross months in 2013 leads to $M_{11,2013M12,jan} = 0.0010$ since only 0.10% of interviews took place in January, $M_{11,2013M12,feb} = 0.2656$ since 26.56% of the sample were interviewed in February, $M_{11,2013M12,mar} = 0.2824$ for a proportion of 28.24% of interviews in March, and so on. In contrast, $M_{11,2013M01,m}, \ldots, M_{11,2013M11,m} = 0$ for all m, since the yearly value is brought into the model in the last month of the year, and hence all other months are missing. The Microcensus measures the same underlying θ_{1t} as the GSOEP, but by quarterly averages according to (7), so that $M_{21,t}(L) = \frac{1}{3} + \frac{1}{3}L + \frac{1}{3}L^2$ for each t reflecting the last month of the quarter (March, June, September, or December), and $M_{21,t}(L) = 0$ for other values of t. The Ifo measure of overtime refers to a single month, and hence $M_{33,t}(L) = 1$ if data are available in month t and $M_{33,t}(L) = 0$, otherwise.

Since for a long sample of data prior to 2010 the GSOEP is the only available statistic directly measuring θ_{1t}, we implement this source as a benchmark, and force a weighted average of θ_{1t} to fit the yearly GSOEP figure exactly. Recent Microcensus figures, in contrast, enter the model with an adjustment term d_2, and the survey error ε_{2t} is modeled with a fixed variance which is estimated within the state space model. Since quantitative information on the survey autocorrelation due to overlap is not available, we model the survey error as serially uncorrelated.

Since the Ifo series does not directly measure the overtime hours, but rather serve as a correlated indicator, the inclusion of a measurement error is not important here. The measurement error would be anyway exchangeable with the noise term in the dynamic model, because the Ifo series estimates a single underlying series, and because the measurement scheme does not involve a filter of lagged values so that there is no autocorrelation induced by the measurement scheme.

We choose an unrestricted multivariate STSM formulation of θ_t^C and θ_t as the dynamic model. In contrast to the Formulation (3), this approach allows for correlation between θ_{1t} and θ_{2t} (fraction of establishments with overtime) beyond their dependence on the common components, while nesting the strict factor specification. The need for this additional flexibility is reasonable since the Ifo survey measures a concept relatively close to the target series, and may provide specific information beyond the overall business cycle.

A large gain in parsimony is associated by using only one factor instead of all ten indicators in the model and hence drastically reducing the model by means of reduced rank sparsity. Additionally, however, we conduct model selection and reduce the model parameters by dropping different individual components from the model. The decision to drop components is drawn from sequential tests based on each series individually. Augmented Dickey Fuller tests, with lag lengths determined by AIC, fail to reject unit roots for each of the series considered in the models (the exception being $\bar{f}_{1t}$, with a p-value of 0.0029). We hence include unit root components for all series and let $\xi_{it} \neq 0$ in general. We test the presence of slope changes ζ_t, white noise terms u_t, cyclical components c_t, and changes to the seasonal pattern ω_t in univariate STSMs, and present the p-value of the corresponding hypotheses in Table 5. The time series of Microcensus data is not sufficiently long for univariate analyses so that we base the specification for paid and unpaid overtime on the yearly GSOEP series.

On a five percent significance level, there is no evidence for a changing slope in either of the series. This again supports the computation of principal components based on first differences rather than second differences of the data. We hence set Σ_ζ to zero in both the model for paid and for unpaid overtime hours. According to additional test results, we include a white noise term for the principal components, but not for the series in z_t in what follows. There is relatively strong evidence on the presence of cyclical components which seems to be needed in each of the observed series. Finally, the Ifo survey is the only series with a seasonal component that is reasonably modeled with a fixed seasonal pattern. The

Table 5. P-values from testing different null hypotheses on the presence of several components in univariate structural time series models. The tests refer to the full model in the alternative. Models are formulated at the original data frequency (monthly for $\hat{f}_t$ yearly for GSOEP, quarterly for Ifo Business Survey).

	$\hat{f}_{1t}$	$\hat{f}_{2t}$	Paid Ot. (GSOEP)	Unpaid Ot. (GSOEP)	Overtime (Ifo)
$H_0 : \varsigma_t = 0$	0.4213	0.0898	0.6815	0.9735	1.0000
$H_0 : u_t = 0$	0.0008	0.0433	1.0000	0.0676	1.0000
$H_0 : c_t = 0$	0.0000	0.0010	0.0053	0.0133	0.0000
$H_0 : \omega_t = 0$	—	—	—	—	0.4396

Table 6. Estimated parameters for cyclical components in models for paid and unpaid overtime hours (first two columns) and flows on working time accounts (last two columns). A similar cycles assumption is imposed in each of the models.

	Paid Ot.	Unpaid Ot.	Inflow WTA	Outflow WTA
Dampening factor ρ	0.9832	0.9880	0.9835	0.9835
Angle frequency λ	0.1155	0.1128	0.1198	0.1198
Period $\frac{2\pi}{\lambda}$	54.42	55.70	52.45	52.45
Cycle standard deviation	8.45	3.89	15.04	9.34
Cycle shock correlation with κ_t^c	0.69	0.45	0.60	-0.42
Phase shift δ	-2.60	-3.15	0.89	-7.79

seasonal figure in overtime hours comes in only trough the short Microcensus time series and has therefore also to be set fixed.

Considering the joint dynamic process introduced above, a decomposition

$$\begin{pmatrix} \theta_t^C \\ \theta_{1t} \\ \theta_{2t} \end{pmatrix} = \begin{pmatrix} \mu_t^C \\ \mu_{1t} \\ \mu_{2t} \end{pmatrix} + \begin{pmatrix} 0 \\ \gamma_{1t} \\ \gamma_{2t} \end{pmatrix} + \begin{pmatrix} c_t^C \\ c_{1t} \\ c_{2t} \end{pmatrix} + \begin{pmatrix} u_t^C \\ 0 \\ 0 \end{pmatrix}$$

applies, with dynamic components driven by the processes introduced below Equation (1). There, Σ_κ and Σ_ξ are full symmetric $(r+2)\times(r+2)$ parameter matrices, while Σ_u is a scalar and $\Sigma_\zeta = \Sigma_\omega = 0$.

Both for paid and unpaid overtime hours, models with $r = 1$ are estimated as the baseline specifications, which appears reasonable due to the relatively small number of indicators in x_t and avoids parameter abundance. Setting $r = 2$ while using the same modeling strategy does not change the estimated time series in a relevant way. We assess whether the data are consistent with a similar cycles assumption ($\rho_i = \rho$, $\lambda_i = \lambda$) and whether the overtime measures θ_{1t} and θ_{2t} have the same cycle shift with respect to the business cycle factor ($\delta_2 = \delta_3$). These restrictions are rejected neither for paid, nor for unpaid overtime hours on a 5% significance level, so that they are maintained. The estimated cyclical parameters are shown in the left two columns of Table 6.

For both models, we find that the cycles are relatively persistent, with a dampening factor close to one, and that a typical cycle lasts about four and a half years. The cycles are shifted by approximately three months to the right relative to the business cycle of the principal component, so that a peak in overtime hours typically lags behind that of the factor. Paid overtime hours appear to be more pro-cyclical, since the standard deviation of the factor (log-scale $\times$ 100) is more than twice as large as that for unpaid overtime hours. At the same time, paid overtime hours exert a stronger correlation with the business cycle.

Further results on the volatility and correlations of cycle and trend shocks are given for paid overtime in Table 7 and for unpaid overtime in Table 8. We observe strong positive cycle correlations also for the Ifo overtime hours, which justifies the inclusion of this series. From the standard deviations of ξ_{1t} for both model, it can be seen that the trend is

Table 7. Estimated standard deviations (main diagonal) and correlations (below diagonal) of cycle shocks κ_t (left) and trend shocks ξ_t (right) for paid overtime model.

	κ^C	κ_1	κ_2		ξ^C	ξ_1	ξ_2
κ^C	1.10	—	—	ξ^C	0.60	—	—
κ_1	0.69	1.55	—	ξ_1	− 0.70	1.72	—
κ_2	0.95	0.43	5.78	ξ_2	− 0.94	0.41	4.12

more volatile for unpaid than for paid overtime, mirroring the larger impact of the cycle on the paid overtime hours.

Figure 2 shows the smoothed estimate for paid overtime hours. The observations of the GSOEP (round points, placed in March of each year) and Microcensus (crosses, net of the constant d_2) are shown along with the trend μ_{1t} (dotted), the seasonally adjusted estimate $\mu_{1t} + c_{1t}$ (dashed) and the overall smoothed series including the seasonal component (solid). The ordinate axis is depicted on a logarithmic scale to reflect the logarithmic model formulation. The nearly linear long-term downward trend is visibly superimposed by stochastic cycles which had a pronounced effect during the 2008/09 financial and economic crisis and reflects well-known patterns from cyclical output movements. The fixed seasonal component, which shows higher overtime usage in the second half of the year, stems mostly from the short sample of Microcensus observations, and should therefore be treated with care.

The unpaid overtime hours, shown in Figure 3, are driven by a rather volatile trend which closely follows the observations. There are several periods of longer upward or downward movements, and although unpaid hours rose in tendency over the whole sample, there is a decline since about 2006 until now. The cycle is rather small, which reflects the low business cycle sensitivity of this working-time component, while the seasonal component is positive in the first and fourth quarter.

We assess the stability of the models by studying estimated parameters in different subsamples. We estimate both the paid and the unpaid overtime models for subsamples of two thirds of the monthly observations (200 of the 300 months from 1991 to 2015). Table 9 shows results for paid overtime in the left and results for unpaid overtime in the right panel, where the subsample ranging from January 1991 to August 2007 is denoted by Smpl 1, the subsample from March 1995 to October 2011 is denoted by Smpl 2, and the subsample from May 1999 to December 2015 is denoted by Smpl 3.

We find notable differences in the cyclical properties: The first subsample has a more persistent cycle (higher ρ), smaller frequency (smaller λ) and a longer phase shift of the overtime variables (absolutely larger ξ) for both models. Also the standard errors of trend

Table 8. Estimated standard deviations (main diagonal) and correlations (below diagonal) of cycle shocks κ_t (left) and trend shocks ξ_t (right) for unpaid overtime model.

	κ^C	κ_1	κ_2		ξ^C	ξ_1	ξ_2
κ^C	1.14	—	—	ξ^C	0.41	—	—
κ_1	0.45	0.60	—	ξ_1	0.21	2.87	—
κ_2	0.94	0.74	5.51	ξ_2	− 1.00	− 0.28	4.54

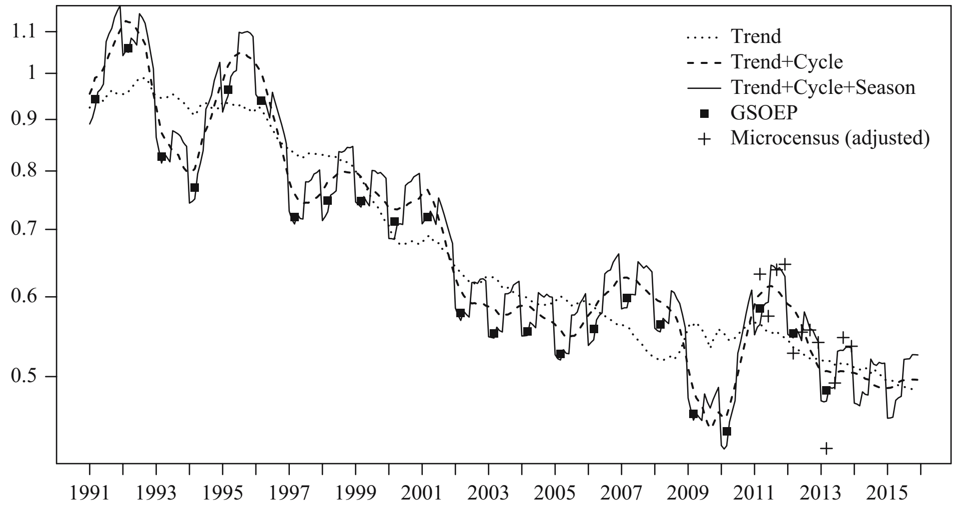

Fig. 2. Paid overtime hours per week. The trend, cycle and seasonal figures are obtained by the state smoother and shown along with the GSOEP and Microcensus observations. The latter is adjusted for the constant d_2.

shocks ($sd(\xi_{jt})$) are subject to change, most prominently for the first and second series of the paid overtime model, where the standard deviations change by a factor of two or more, and large trend variance in a given sample is associated with a smaller cycle variance (smaller $sd(\kappa_{jt})$). The correlations ($corr(\xi_{jt}, \xi_{it})$ and $corr(\kappa_{jt}, \kappa_{it})$) even change sign in some cases. The apparent structural instability is due to multiple maxima of the likelihood function, where different local maxima dominate in different subsamples. Improved stability, for example, by averaging over different local maxima or applying numerical

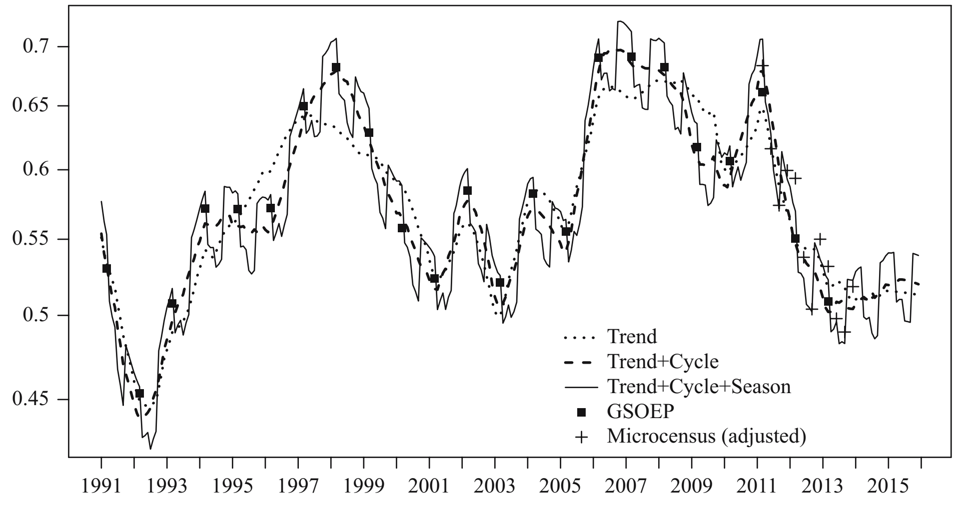

Fig. 3. Unpaid overtime hours per week. The trend, cycle and seasonal figures are obtained by the state smoother and shown along with the GSOEP and Microcensus observations. The latter is adjusted for the constant d_2.

Table 9. Estimated parameters for paid (left) and unpaid (right) overtime models and different subsamples each with two-thirds of the observations: First (Smpl 1), middle (Smpl 2), and last portion of the data (Smpl 3).

Par.	Smpl 1	Smpl 2	Smpl 3
ρ_1	0.9926	0.9803	0.9805
ρ_2	0.9926	0.9803	0.9805
ρ_3	0.9926	0.9803	0.9805
λ_1	0.0714	0.1064	0.0995
λ_2	0.0714	0.1064	0.0995
λ_3	0.0714	0.1064	0.0995
ξ_1	0.0000	0.0000	0.0000
ξ_2	−9.4151	−2.9368	−4.8008
ξ_3	−9.4151	−2.9368	−4.8008
$sd(\xi_{1t})$	0.3430	0.1706	0.5064
$sd(\xi_{2t})$	1.0588	2.5260	1.4812
$sd(\xi_{3t})$	5.6845	5.9511	5.7918
$corr(\xi_{1t}, \xi_{2t})$	−0.9936	−0.7646	−0.9998
$corr(\xi_{1t}, \xi_{3t})$	−0.9998	−0.3016	0.3731
$corr(\xi_{2t}, \xi_{3t})$	0.9956	−0.3837	−0.3893
$sd(\kappa_{1t})$	0.9774	1.3254	1.2326
$sd(\kappa_{2t})$	2.6402	0.8432	1.3161
$sd(\kappa_{3t})$	6.9054	5.9609	5.4484
$corr(\kappa_{1t}, \kappa_{2t})$	0.7252	0.9432	0.9397
$corr(\kappa_{1t}, \kappa_{3t})$	0.9372	0.8800	0.8430
$corr(\kappa_{2t}, \kappa_{3t})$	0.4404	0.6725	0.9761

Par.	Smpl 1	Smpl 2	Smpl 3
ρ_1	0.9944	0.9900	0.9809
ρ_2	0.9944	0.9900	0.9809
ρ_3	0.9944	0.9900	0.9809
λ_1	0.0718	0.1065	0.0999
λ_2	0.0718	0.1065	0.0999
λ_3	0.0718	0.1065	0.0999
ξ_1	0.0000	0.0000	0.0000
ξ_2	−6.9145	−5.3627	−2.9933
ξ_3	−6.9145	−5.3627	−2.9933
$sd(\xi_{1t})$	0.3935	0.6908	0.2875
$sd(\xi_{2t})$	3.6219	3.1814	3.0331
$sd(\xi_{3t})$	5.7547	6.1983	5.5740
$corr(\xi_{1t}, \xi_{2t})$	0.4744	0.9061	0.7796
$corr(\xi_{1t}, \xi_{3t})$	−0.9431	0.5471	0.1223
$corr(\xi_{2t}, \xi_{3t})$	−0.1546	0.1417	−0.5262
$sd(\kappa_{1t})$	0.9394	1.1645	1.2796
$sd(\kappa_{2t})$	0.1545	0.6507	0.7645
$sd(\kappa_{3t})$	6.8822	5.8553	5.3710
$corr(\kappa_{1t}, \kappa_{2t})$	−0.9954	−0.5995	0.4459
$corr(\kappa_{1t}, \kappa_{3t})$	0.9998	0.8345	0.8958
$corr(\kappa_{2t}, \kappa_{3t})$	−0.9963	−0.0594	0.7973

integration in a Bayesian approach are beyond the scope of this article, but are an important topic of future research.

To illustrate the effect of using a multivariate approach, we contrast the result to a univariate structural time series approach. To mimic the dynamic specification in the multivariate approach, we fit a model consisting of level and cycle to the yearly GSOEP series, but omit the seasonal which would not be identified. We use the same measurement scheme as $M_{11,t}(L)$ above in the multivariate model. In Figure 4 the seasonally adjusted multivariate estimate (dashed line) for paid overtime hours is shown along with the GSOEP observations (points, again placed in March of each year) and the smoothed estimate from the univariate approach (solid line). We see only slight differences between the lines in the time span before 2013: The GSOEP observations are used as a benchmark for the yearly weighted means of the estimates, and hence the latter do not move too far away from the former. As a slight deviation, we see that a decline in overtime hours before the global financial crisis is detected already in 2007 using all available indicators, while the univariate GSOEP model gives a smoother change to the "crisis regime". After 2010 when the Microcensus data come in and especially after the last GSOEP observation the multivariate estimates clearly uses more relevant information than just an extrapolation of the dynamics. Hence the additional indicators have the greatest impact where information is most valuable: at the current edge. The figure for unpaid overtime hours, which shows only minor differences between the univariate and the multivariate model, is available from the author upon request.

Finally, we investigate the advantage of our modeling approach with respect to the quality of early estimates in official statistics. A numerical assessment of the overall precision of our estimates of paid and unpaid overtime hours is not possible: The truth is not known and hence a straightforward benchmark as in the simulation study is not

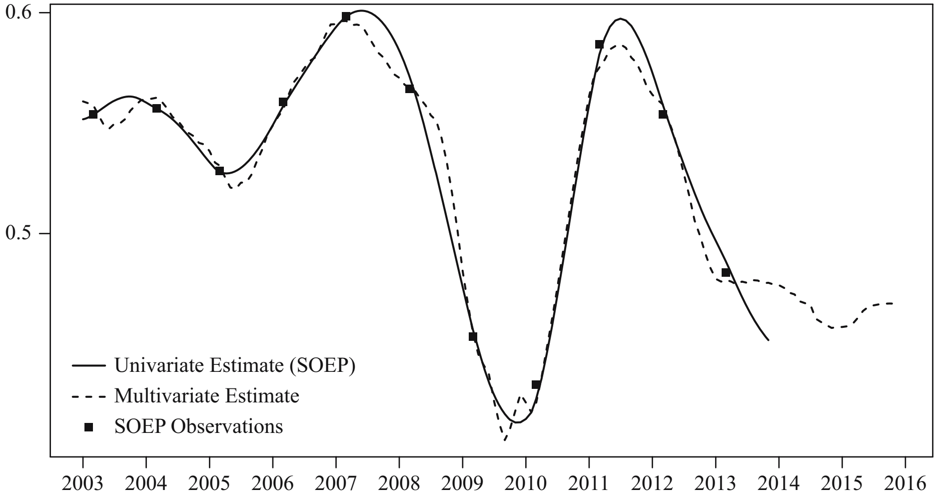

Fig. 4. Paid overtime hours per week. The seasonal adjusted multivariate estimate (dashed) is shown along with a univariate structural time series model (solid) using SOEP data alone. The univariate model implies the same dynamic properties as the multivariate model described in Subsection 5.1 and uses the same measurement scheme with respect to the SOEP observations.

available. We can however assess the timeliness of precise estimates by an ex-post comparison of early estimates in real time to the final estimates where all data are available. We figure out the value of additional information beyond the GSOEP data by comparing the multivariate model to the univariate benchmark. Since the seasonality in the multivariate model can be estimated only in recent years, the seasonal patterns in the Microcensus would cause problems in the pseudo real-time study. We hence omit the Microcensus data from the multivariate model and ignore seasonality both for the univariate and multivariate model in our study.

We conduct the real-time experiment as follows: For the construction of data for a given quarter, we take the indicators contained in x_t and the Ifo Business survey for that quarter as known. The GSOEP data are used with a time lag; the data from two years earlier are used so that for example from the beginning of 2015 the data for 2013 are available. This is a realistic timing since first estimates are typically constructed in the middle of the subsequent quarter. Having constructed pseudo-real-time estimates by a Kalman filter for each of the two models and each quarter from 2004Q1 to 2015Q4, we compare them to the ex-post estimate of each model using all available data. From this comparison, the biases and root mean squared errors can be computed which are shown in Table 10. There, along with the full evaluation sample (all), also results for successive subsamples are given which divide the evaluation sample in four intervals of three years.

Overall, in terms of the RMSE, we find that the multivariate model outperform the univariate approach for paid overtime, while the approaches perform similar for unpaid overtime where the univariate approach is slightly better. This is in line with the finding that paid overtime is more correlated to business cycle indicators and thus the latter help estimate paid overtime hours better than unpaid hours. The very parsimonious univariate model has a smaller bias for both target measures. Considering the subsamples, it is reassuring for the multivariate approach that the latter outperforms the univariate approach in the last subsample, where more data are used in the Kalman filter. The multivariate model is likely to dominate in larger samples which is reflected here. In sum, it is preferable to use a factor approach as the one considered in this article if the target series has a strong business cycle correlation, and if long enough time series are available.

Table 10. Bias and root mean squared error (RMSE) for paid (left) and unpaid (right) overtime model and univariate (univ.) and multivariate (mult.) model. An evaluation sample from 2003Q4 to 2015Q4 is used (all) and four successive subsamples thereof.

	Paid Overtime				Unpaid Overtime			
	BIAS		RMSE		BIAS		RMSE	
Smpl	univ.	mult.	univ.	mult.	univ.	mult.	univ.	mult.
all	−4.047	−5.680	10.846	9.582	0.217	2.373	11.064	11.402
1	4.212	−0.761	7.681	8.126	−2.888	−11.052	10.648	12.830
2	−11.157	−12.923	14.241	13.613	−4.426	2.935	11.991	9.268
3	−4.712	−9.152	13.034	10.665	−0.059	10.311	8.521	11.451
4	−4.530	0.116	6.232	1.474	8.243	7.300	12.643	11.764

5.2. Net Flows on Working Time Accounts

Not all additional hours worked by employees in a given period lead to a definitive increase in the amount of labor over a longer time span. Some of them, termed transitory overtime hours, are compensated by leisure time in a future period. The number of these additional hours worked hence raise the credits on WTA, which are formal arrangements to record such additional hours worked. When measuring hours actually worked per period, the statistician has to track such inflows on WTA which raise hours worked, but also the outflows from WTA which reduce the overall hours worked.

Only few data sources are available which allow to measure in- and outflows from WTA in Germany on a regular basis. Besides paid and unpaid overtime hours, the GSOEP questionnaire asks for overtime hours which are compensated with time-off, and which we hence treat as inflows on WTA. A question regarding the reduction of such hours has been included in the questionnaire only in 2014 and the results are not yet available. A similar objection is faced by a new question regarding balances on WTA in the IAB Job Vacancy Survey. It has been included in the establishment survey in 2013 and therefore still lacks a sufficient history to base long time series estimates thereupon.

The Microcensus holds additional information on WTA flows over a longer time span, which we exploit in our estimation strategy. Each employed household member is asked for the regular weekly hours worked and for hours worked last week. If both differ, the main reason for that difference is inquired, where possible answers include "compensation for more hours worked (e.g., flexible working hours)" if actual hours were lower and "hours for the accumulation of the time credit or for the reduction of time dept" if they were higher than usual. These or analogous questions are available for the whole estimation period.

Since only the *main* reason for a difference is asked for in the Microcensus, there are likely further WTA in- or outflows that are not revealed by the survey participants and hence the results are biased. Our strategy thus combines information on the level of gross inflows from the GSOEP with cyclical variations of the Microcensus figures on in- and outflows around their trends to arrive at a final estimate of net flows. The maintained assumption is that even if both WTA in- and outflows follow (possibly stochastic) trends, the latter should be identical so that there is no long-run discrepancy between the both, and the net flows average to zero in the long run. This allows us to estimate the trend by use of the GSOEP series, while relative deviations from it are determined from the Microcensus. Stated jointly with the estimated factors, the measurement model is

$$\begin{pmatrix} \bar{f}_t \\ \log(in\_mc_t) \\ \log(out\_mc_t) \\ \log(in\_gsoep_t) \end{pmatrix} = \begin{pmatrix} 0 \\ d_{1t} \\ d_{2t} \\ 0 \end{pmatrix} + \begin{pmatrix} I & 0 & 0 & 0 \\ 0 & M_{11,t}(L) & 0 & 0 \\ 0 & 0 & M_{22,t}(L) & 0 \\ 0 & 0 & 0 & M_{33,t}(L) \end{pmatrix} \begin{pmatrix} \theta_t^C \\ \theta_{1t} \\ \theta_{2t} \\ \theta_{3t} \end{pmatrix}.$$

The measurement polynomials $M_{11,t}(L)$, $M_{22,t}(L)$ and $M_{33,t}(L)$ are again designed to fit the characteristics of the Microcensus and GSOEP surveys and in particular the distribution of interviews over the time-spans for which the surveys can be distinctly evaluated. In contrast to the overtime models in Subsection 5.1, for which Microcensus data are available only since 2010, the restructuring of the Microcensus has to be taken into account in the current WTA model. In 2005, a fixed reference week each year (or less frequently before 1995) was replaced by a continuous interviewing policy. This leads to a change in the polynomials, which are given by $M_{11,t}(L) = M_{22,t}(L) = 1$ before 2005 and $M_{11,t}(L) = M_{22,t}(L) == \frac{1}{3} + \frac{1}{3}L + \frac{1}{3}L^2$ since then. Additionally, a level shift results from the changing survey practice at this time which we model by setting d_{1t} and d_{2t} to nonzero constants before 2005 and to zero afterwards. For the GSOEP measurement, given by $M_{33,t}(L)$, we use the same approach as for the paid and unpaid overtime models described in Subsection 5.1.

In contrast to the case where two surveys measure the same underlying process, in the case of one survey per series the survey error variance cannot be estimated from the data when the underlying process has an additional noise term. The survey error variance could be rather set fixed, based on further information from the survey design. As the pure sampling uncertainty is very small for the large sample of the Microcensus, we do not model survey errors explicitly in this case and set $\varepsilon_t = 0$, while allowing irregular components within the model for θ_t.

We again conduct model selection by studying the individual processes first, and include only components in the joint model which appear worthwhile from the univariate tests. We thus again include a unit root component for all series in order to reflect results from Augmented Dickey Fuller tests. A univariate analysis of the individual components similar to Table 5 reveals that the GSOEP inflow series has a significant slope change (p-value 0.04), while a noise term finds more support from the data than a cycle (for which the p-value is 0.16).

For the Microcensus series, we set $\zeta_t = 0$ and include a noise term along with the cycle and random walk trends, which is also supported by statistical tests in the multivariate model. As an a-priori modeling decision to gain parsimony, correlations between the GSOEP and other series are not considered and hence the former is used solely to extract its trend by univariate filtering and smoothing. The model is thus given by

$$\begin{pmatrix} \theta_t^C \\ \theta_{1t} \\ \theta_{2t} \\ \theta_{3t} \end{pmatrix} = \begin{pmatrix} \mu_t^C \\ \mu_{1t} \\ \mu_{2t} \\ \mu_{3t} \end{pmatrix} + \begin{pmatrix} 0 \\ \gamma_{1t} \\ \gamma_{2t} \\ 0 \end{pmatrix} + \begin{pmatrix} c_t^C \\ c_{1t} \\ c_{2t} \\ 0 \end{pmatrix} + \begin{pmatrix} u_t^C \\ u_{1t} \\ u_{2t} \\ u_{3t} \end{pmatrix},$$

where Σ_ζ has a single nonzero element associated with θ_{3t}, Σ_u is diagonal, and Σ_ξ as well as Σ_κ are block diagonal with a full upper left 3×3 submatrix.

The properties of the cyclical components of WTA in- and outflows are summarized in the right two columns of Table 6. Again, we cannot reject the similar cycles restriction, and the common period and the dampening factor are similar to the case of overtime hours. Both components have a relatively strong cyclical pattern, and WTA inflows have the highest cycle standard deviation among the variables under consideration. Not surprisingly, shocks to inflows are positively, while outflow shocks are negatively related to business cycle shocks. The phase shifts mean that typically seven months after employees have built up most credit on the accounts, the outflows peak and reduce the savings on WTA.

The cyclical patterns of in- and outflows are shown in Figure 5, where $c_{jt} + u_{jt}, j = 1,2$, is depicted for inflows (solid line) and outflows (dashed line) in logarithmic scale $\times$ 100, as annotated on the left axis. The mentioned phase shift between the cycles becomes evident here. At most of the visible peaks of WTA inflows, the outflows are rising and reach their highest value a few months later. Before the building up of credits beginning in 2005, the outflows dropped, while the credits were used up afterwards during the 2008/09 crisis, where outflows peaked again. The trending behavior of transitory overtime hours from the GSOEP, which is used as the trend in both, in- and outflows from WTA, is shown in hours per week as the thin dash-dotted line with annotation at the right axis. It shows a flattening growth from below 0.5 hours per week to over one hour until 2010, and has diminished slightly over the recent years.

The trend and log-scale cycles are combined multiplicatively to yield the net flow on WTA, which is the relevant statistic measuring the effect on hours worked per period. We compute this effect as

$$\Delta WTA_t \approx \exp\left(\mu_{3t}^z\right)\left(\gamma_{1t}^z + c_{1t}^z + u_{1t}^z - \gamma_{2t}^z - c_{2t}^z - u_{2t}^z\right). \tag{9}$$

This overall effect is plotted in Figure 6, where also the seasonal patterns are assessed. The overall increase in the scale of the fluctuations over time is partly due to the increased overall importance of WTA corresponding to the upward trend of gross flows described above, while the cyclical patterns from Figure 5 are closely reflected by the overall net flows.

As for the results of paid and unpaid overtime hours, further processing of the data is performed within the working-time measurement concept to yield quarterly results which are partly decomposed for several groups of employees. These are published by the IAB in the form of working time components tables, and also enter the publication of national accounts by the German Federal Statistical Office.

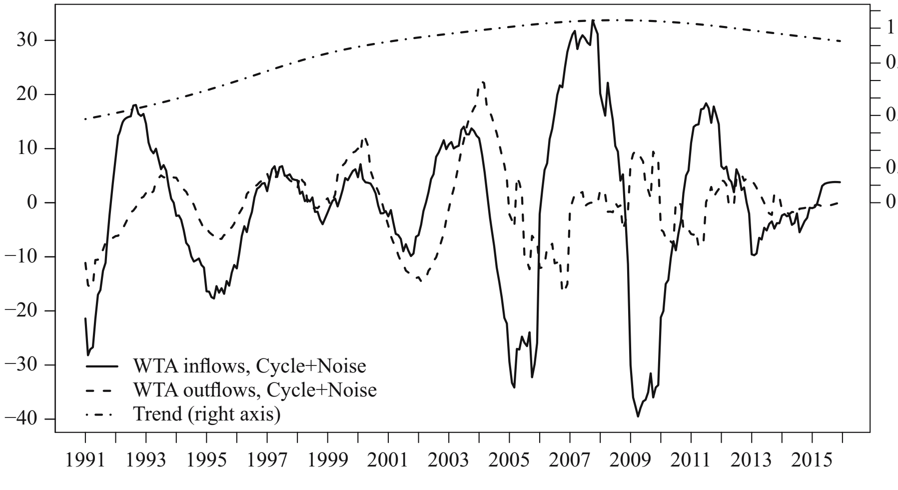

Fig. 5. Cyclical and noise components of in- and outflows (left axis) and trend in flows on working time accounts (right axis). The cycle, noise and trend figures are obtained by the state smoother. Cycles and trends are combined multiplicatively according to (9) to obtain estimated WTA net flows.

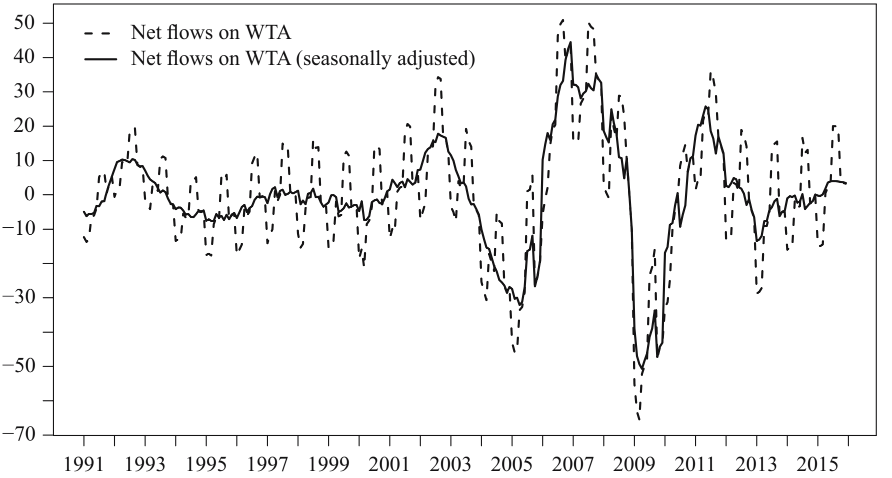

Fig. 6. Working time account net flows in hours per week, computed from smoothed cycles and trends by $\Delta WTA_t \approx \exp\left(\mu_{3t}^{z}\right)\left(\gamma_{1t}^{z} + c_{1t}^{z} + u_{1t}^{z} - \gamma_{2t}^{z} - c_{2t}^{z} - u_{2t}^{z}\right)$.

6. Conclusion

We have proposed a factor structural time series model and discussed its implementation for possibly high-dimensional problems in official statistics. The model has an intuitive appeal due to its additive, componentwise structure and is quite unrestrictive in its formulation. It is straightforward to apply using a two-step approach with principal components and state-space techniques. In simulation experiments, we found that the two-step approach works reasonably well and that the new method outperforms several competitors in terms of its ability to estimate an unobserved target series. These results show the potential to construct more timely and precise official statistics that use a wide array of recently available data. The empirical application in the article illustrated the usefulness of the method for the measurement of working time components. There, the model was used to construct a time series that is longer, more frequent, and uses more recent information than single survey data sources alone.

The main motivation of the approach was for smoothing latent series using surveys and several other indicators, as is of foremost importance for statistical agencies. However, the methods may reveal their strength also for other tasks such as exploration of the componentwise dynamic properties and co-movements of several macroeconomic time series, as well as forecasting. Additional research may also be concerned with correlated unobserved components models in high dimensions, which allow for a more flexible modeling of spillovers and structural identification.

Appendix A

Properties of the Factor Model in Differences

In this appendix, we show that the data generated by a factor STSM satisfy strong assumptions on time and cross-section dependence when suitably differenced. These are sufficient to ensure key results of Bai and Ng (2002). Denoting $X_{it} := \Delta\Delta_s y_{it}$, $i = 1, \ldots, N$, and $F_{jt} := \Delta\Delta_s f_{jt}$, $j = 1, \ldots, r$, where y_t is the vector of observed variables and f_t is the compound factor process introduced in the main text, the factor model can be stated as

$$X_{it} = \Lambda_{i}.F_t + e_{it}.$$

Here, e_{it} is cross-sectionally uncorrelated and independent from F_t by assumption, while both F_{jt} and e_{it} follow strictly stationary linear Gaussian processes with absolutely summable coefficients, as we discuss in the following.

To see the dynamic properties more clearly, a generic element from e_{it} is stated as

$$e_{it} = \Delta\Delta_s\mu_{it} + \Delta\Delta_s\gamma_{it} + \Delta\Delta_s c_{it} + \Delta\Delta_s u_{it},$$

where the I superscript is suppressed for notational simplicity. Since $\mu_{it} = \mu_{i,t-1} + \nu_{i,t-1} + \xi_{i,t-1}$, we have $\Delta\mu_{it} = \nu_{i,t-1} + \xi_{i,t-1}$, while from $\nu_{it} = \nu_{i,t-1} + \zeta_{i,t-1}$, it follows that $\nu_{it} = \nu_{i,t-s} + \zeta_{i,t-1} + \ldots + \zeta_{i,t-s}$. Hence,

$$\Delta\Delta_s\mu_{it} = \Delta_s\nu_{i,t-1} + \Delta_s\xi_{i,t-1} = \zeta_{i,t-2} + \ldots + \zeta_{i,t-s-1} + \xi_{i,t-1} - \xi_{i,t-s-1},$$

where ξ_{it} and ζ_{it} are mutually independent Gaussian iid processes, and hence a finite-order moving average structure is obtained for the differenced trend component, with coefficients straightforwardly obtained from the s nonzero autocovariances.

A similar result is obtained for the seasonal component $\gamma_{it} = -\gamma_{i,t-1} - \ldots - \gamma_{i,t-s+1} + \omega_{i,t-1}$. Applying first differences to both sides of this equation yields $\gamma_{it} = \gamma_{i,t-s} + \omega_{i,t-1} - \omega_{i,t-2}$, and hence

$$\Delta\Delta_s\gamma_{it} = \Delta^2\omega_{i,t-1},$$

which is again a (over-differenced) finite-order moving average that trivially has absolutely summable coefficients.

Regarding the cycle, Harvey (1991, Sec. 2.5.6) gives the stationary ARMA(2,1) representation for $|\rho_i| < 1$, which leads directly to

$$\Delta\Delta_s c_{it} = \Delta\Delta_s \frac{1 + \theta_i L}{1 - 2\rho_i \cos(\lambda_i) L - \rho_i^2 L^2} \tilde{\kappa}_{i,t-1},$$

where θ_i is a moving average parameter and $\tilde{\kappa}_{it}$ is composed of the two jointly Gaussian iid processes κ_{it} and κ^*_{it}. Since as a stationary ARMA process the fraction expands to a polynomial with absolutely summable coefficients, also the entire expression for $\Delta\Delta_s c_{it}$ shares this property while inheriting stationarity and Gaussianity. The same is true for differenced noise term $\Delta_4\Delta u_{it}$. Hence, any linear combination of $\Delta\Delta_s\mu_{it}$, $\Delta\Delta_s\gamma_{it}$, $\Delta\Delta_s c_{it}$ and $\Delta\Delta_s u_{it}$ is strictly stationary, Gaussian and has absolutely summable coefficients. The statement is applicable both to the differenced idiosyncratic components e_{it} and to series of the differenced factor process F_t.

The properties of F_t and e_{it} are clearly sufficient to assure Assumptions A (by a law of large numbers drawing on ergodicity of F_t), C (since absolutely summable autocovariances follow from absolutely summable Wold coefficients), and D (due to the independence between e_{it} and F_t) of Bai and Ng (2002), while their Assumption B on the factor loadings has to be imposed additionally to obtain the main results of that article.

Clearly, the squared autocorrelations of e_{it} are also summablein our setup, and hence Bai and Ng (2002, eq. (6)) yields mean-square convergence of estimated F_t to the true values for a given t. Naturally, the consistency holds also for a cumulation of finitely many estimated F_s, $s \leq t$. Hence, also thefactors $\bar{f}_t$ in level are found consistent for a fixed t. The effects of the initial values are lost due to the differencing, however.

Appendix B

The State Space Form

The model given by (3) with measurement scheme (4) can be easily represented in linear state space form which allows to use the techniques described in Durbin and Koopman (2012). We adopt their notation as far as possible and state the system as

$$\begin{pmatrix} \bar{f}_t \\ z_t \end{pmatrix} = z_t \alpha_t + \begin{pmatrix} 0 \\ \varepsilon_t \end{pmatrix}, \quad \varepsilon_t \sim N(0, H_t), \tag{10}$$

$$\alpha_{t+1} = T\alpha_t + R\eta_t, \quad \eta_t \sim N(0, Q), \quad t = 1, \ldots, n. \tag{11}$$

For simplicity of exposition we assume that $l \geq s - 1$, so that l lags of all components have to be included in the state vector to make the measurement equation (4) representable in state space form. Hence, the state vector α_t holds the components μ^I_{it}, μ^C_{jt}, ν^I_{it}, ν^C_{jt}, γ^I_{it}, γ^C_{jt}, $(\tilde{c}^I_{it}, \tilde{c}^{I,*}_{it})$, $(\tilde{c}^C_{jt}, \tilde{c}^{C,*}_{jt})$, u^I_{it} and u^C_{jt}, each for $i = 1, \ldots, N_z$ and $j = 1, \ldots, r$, along with l lags of each component. More precisely,

$$\begin{aligned}
\alpha'_t = (&\mu^I_{1t}, \ldots, \mu^I_{N_z,t}, && \mu^C_{1t}, \ldots, \mu^C_{r,t}, \\
&\nu^I_{1t}, \ldots, \nu^I_{N_z,t}, && \nu^C_{1t}, \ldots, \nu^C_{r,t}, \\
&\gamma^I_{1t}, \ldots, \gamma^I_{N_z,t}, && \gamma^C_{1t}, \ldots, \gamma^C_{r,t}, \\
&\tilde{c}^I_{1t}, \tilde{c}^{I*}_{1t}, \ldots, \tilde{c}^I_{N_z,t}, \tilde{c}^{I*}_{N_z,t}, && \tilde{c}^C_{1t}, \tilde{c}^{C*}_{1t}, \ldots, \tilde{c}^C_{r,t}, \tilde{c}^{C*}_{r,t}, \\
&u^I_{1t}, \ldots, u^I_{N_z,t}, && u^C_{1t}, \ldots, u^C_{r,t}, \\
&\mu^I_{1,t-1}, \mu^I_{2,t-1}, && \ldots \text{lagged components} \ldots, u^C_{r,t-l})'
\end{aligned}$$

is the $m := 6(N_z + r)(l + 1)$-dimensional state vector. Accordingly, the $6(N_z + r) \times (l + 1) \times 6(N_z + r)(l + 1)$ transition matrix is given by

$$T = \begin{pmatrix} \tilde{T} & 0 & \ldots & 0 \\ I & & & \vdots \\ & \ddots & & \vdots \\ 0 & & I & 0 \end{pmatrix}, \quad \text{where} \quad \tilde{T} = \begin{pmatrix} T_\mu & T_{\mu\nu} & 0 & 0 & 0 \\ 0 & T_\nu & 0 & 0 & 0 \\ 0 & 0 & T_\gamma & 0 & 0 \\ 0 & 0 & 0 & T_c & 0 \\ 0 & 0 & 0 & 0 & T_u \end{pmatrix}$$

is a $6(N_z+r)\times 6(N_z+r)$ matrix with $T_\mu = T_\nu = T_{\mu\nu} = I_{N_z+r}$. Moreover, $T_u = 0_{N_z+r}$ and T_c is a $2(N_z+r)\times 2(N_z+r)$ block diagonal matrix with ith block given by

$$T_c^{(i,i)} = \rho_i \begin{pmatrix} \cos\lambda_i & \sin\lambda_i \\ -\sin\lambda_i & \cos\lambda_i \end{pmatrix},$$

for $i = 1,\ldots,N_z+r$. Here, ρ_i and λ_i correspond to the individual cycle parameters for $i = 1,\ldots,N_z$, while they correspond to the parameters of the joint cycles, $\rho_i = \rho^C_{i-N_z}$ and $\lambda_i = \lambda^C_{i-N_z}$ for $i = N_z+1,\ldots,N_z+r$. The transition innovation covariance matrix Q is block diagonal with block element given by Σ^I_ξ, Σ^C_ξ, Σ^I_ζ, Σ^C_ζ, Σ^I_ω, Σ^C_ω, $\Sigma^I_\kappa\otimes I_2$, $\Sigma^C_\kappa\otimes I_2$, Σ^I_u and Σ^C_u, respectively, while R is a vertical stacking of an identity and l quadratic zero matrices that selects the contemporaneous states.

The observation matrices Z_t reflect both the observation patterns for the variables and the loading of common components on the individual series. We denote

$$\tilde{\Lambda} = \begin{pmatrix} 0 & \Gamma_\mu & 0 & 0 & 0 & \Gamma_\gamma & 0 & \check{\Gamma}_c & 0 & \Gamma_u \\ I & \Lambda_u & 0 & 0 & I & \Lambda_\mu & \check{I} & \check{\Lambda}_c & I & \Lambda_\mu \end{pmatrix},$$

where the checked matrices reflect the phase shifts of the variables, so that the ith row of $\check{I}$ is $(\cos(\lambda_i\delta_i), \sin(\lambda_i\delta_i))\otimes I_{i\cdot}$, the ith row of $\check{\Gamma}_c$ is $(\cos(\lambda_i\delta_i), \sin(\lambda_i\delta_i))\otimes \Gamma_{c,i\cdot}$ and the ith row of $\check{\Lambda}_c$ is $(\cos(\lambda_i\delta_i), \sin(\lambda_i\delta_i))\otimes \Lambda_{c,i\cdot}$ which have twice the number of columns as the unchecked quantities. Then, for

$$\begin{aligned}\tilde{M}_t(L) &= \tilde{M}_{t0} + \tilde{M}_{t1}L + \ldots + M_{tl}L^l \\ &= \begin{pmatrix} I & 0 \\ 0 & M_{0t} \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & M_{1t} \end{pmatrix} L + \ldots + \begin{pmatrix} 0 & 0 \\ 0 & M_{lt} \end{pmatrix} L^l\end{aligned}$$

the time-varying observation matrices are given by

$$Z_t = (\tilde{M}_{t0}\tilde{\Lambda}, \tilde{M}_{t1}\tilde{\Lambda}, \ldots, \tilde{M}_{tl}\tilde{\Lambda}),$$

which completes the state space representation for the general case with $d_t = 0$.

If constant terms or statistical breaks occur, the transition matrix is enriched by additional diagonal elements of 1, while the observation matrix reflects this by additional columns with corresponding element either set to the constant values, or switching from zero to that constant at a specified period. The state innovation error covariance matrix is unchanged and the matrix R holds additional rows of zeros.

7. References

Anderson, T.W. 1984. *An Introduction to Multivariate Statistical Analysis*. New York: Wiley.

Bai, J. and S. Ng. 2002. "Determining the Number of Factors in Approximate Factor Models." *Econometrica* 70(1): 191–221. Doi: http://dx.doi.org/10.1111/1468-0262.00273.

Bai, J. and S. Ng. 2004. "A PANIC Attack on Unit Roots and Cointegration." *Econometrica* 72(4): 1127–1177. Available at: http://www.columbia.edu/~sn2294/pub/ecta04.pdf (accessed December 2017).

Bai, J. and S. Ng. 2008. "Large Dimensional Factor Analysis." *Foundations and Trends in Econometrics* 3(2): 89–163. Doi: http://dx.doi.org/10.1561/0800000002.

Bai, J. and S. Ng. 2013. "Principal Components Estimation and Identification of Static Factors." *Journal of Econometrics* 176(1): 18–29. Doi: https://doi.org/10.1016/j.jeconom.2013.03.007.

Banerjee, A., M. Marcellino, and I. Masten. 2014. "Forecasting with Factor-Augmented Error Correction Models." *International Journal of Forecasting* 30(3): 589–612. Doi: https://doi.org/10.1016/j.ijforecast.2013.01.009.

Bernanke, B.S., J. Boivin, and P. Eliasz. 2005. "Measuring the Effects of Monetary Policy: A Factor-Augmented Vector Autoregressive (FAVAR) Approach." *The Quarterly Journal of Economics* 120(1): 387–422. Doi: https://doi.org/10.1162/0033553053327452.

Bollineni-Balabay, O., J. van den Brakel, and F. Palm. 2015. "Multivariate State Space Approach to Variance Reduction in Series with Level and Variance Breaks Due to Survey Redesigns." *Journal of the Royal Statistical Society:* Series A (Statistics in Society). Doi: http://dx.doi.org/10.1111/rssa.12117.

Bräuning, F. and S.J. Koopman. 2014. "Forecasting Macroeconomic Variables Using Collapsed Dynamic Factor Analysis." *International Journal of Forecasting* 30(3): 572–584. Doi: https://doi.org/10.1016/j.ijforecast.2013.03.004.

Breitung, J. and J. Tenhofen. 2011. "GLS Estimation of Dynamic Factor Models." *Journal of the American Statistical Association* 106(495): 1150–1166. Doi: https://doi.org/10.1198/jasa.2011.tm09693.

Burda, M.C. and J. Hunt. 2011. "What Explains the German Labor Market Miracle in the Great Recession?" *Brookings Papers on Economic Activity*, Spring 2011, 273–335. Available at: https://www.brookings.edu/wp-content/uploads/2011/03/2011a_bpea_burda.pdf (accessed December 2017).

Bureau of Economic Analysis. 2017. *Concepts and Methods of the U.S. National Income and Product Accounts*. Technical report. Available at: https://bea.gov/national/pdf/all-chapters.pdf (accessed February 2018).

Commandeur, J., S. Koopman, and M. Ooms. 2011. "Statistical Software for State Space Methods." *Journal of Statistical Software* 41(1): 1–18. Available at: https://www.researchgate.net/publication/260335110_Statistical_Software_for_State_Space_Methods (accessed December 2017).

Croushore, D. 2011. "Frontiers of Real-Time Data Analysis." *Journal of Economic Literature* 49(1): 72–100. Doi: http://dx.doi.org/10.1257/jel.49.1.72.

Durbin, J. 2000. "The State Space Approach to Time Series Analysis and its Potential for Official Statistics." *The Australian and New Zealand Journal of Statistics* 42(1): 1–23. Doi: http://dx.doi.org/10.1111/1467-842X.00104.

Durbin, J. and S.J. Koopman. 2012. *Time Series Analysis by State Space Methods: Second Edition*. Oxford: Oxford Statistical Science Series.

Durbin, J. and B. Quenneville. 1997. "Benchmarking by State Space Models." *International Statistical Review* 65(1): 23–48. Doi: http://dx.doi.org/10.1111/j.1751-5823.1997.tb00366.x.

Eickmeier, S. 2009. "Co-Movements and Heterogeneity in the Euro Area Analyzed in a Nonstationary Dynamic Factor Model." *Journal of Applied Econometrics* 24(6): 933–959. Doi: http://dx.doi.org/10.1002/jae.1068.

Federal Statistical Office. 2008. *National Accounts: Quarterly Calculations of Gross Domestic Product in accordance with ESA 1995 – Methods and Data Sources*. Subject-matter Series 18, S. 23. Available at: https://www.destatis.de/EN/Publications/Specialized/Nationalaccounts/QuarterlyCalculationsGrossDomesticProductAccordance.pdf?__blob=publicationFile (accessed December 2017).

Giannone, D., L. Reichlin, and D. Small. 2008. "Nowcasting: The Real-Time Informational Content of Macroeconomic Data." *Journal of Monetary Economics* 55(4): 665–676. Available at: https://EconPapers.repec.org/RePEc:eee:moneco:v:55:y:2008:i:4:p:665-676 (accessed December 2017).

Harvey, A. and C.-H. Chung. 2000. "Estimating the Underlying Change in Unemployment in the UK." *Journal of the Royal Statistical Society*: Series A (Statistics in Society) 163(3): 303–339. Doi: http://dx.doi.org/10.1111/1467-985X.00171.

Harvey, A. and S. Koopman. 1997. Multivariate Structural Time Series Models. Systematic Dynamics in Economic and Financial Models, (pp. 269–298). Available at: http://personal.vu.nl/s.j.koopman/old/publications/multi_stsm.pdf (accessed February 2018).

Harvey, A.C. 1991. *Forecasting, Structural Time Series Models and the Kalman Filter*. New York: Cambridge University Press.

Helske, J. 2016. *KFAS: Kalman Filter and Smoother for Exponential Family State Space Models*. R package version 1.2.4.

Koopman, S., A. Harvey, J. Doornik, and N. Shephard. 2009. *STAMP 8.2: Structural Time Series Analyser, Modeler, and Predictor*. London: Timberlake Consultants.

Krieg, S. and J.A. van den Brakel. 2012. "Estimation of the Monthly Unemployment Rate for Six Domains through Structural Time Series Modelling with Cointegrated Trends." *Computational Statistics and Data Analysis* 56(10): 2918–2933. Doi: https://doi.org/10.1016/j.csda.2012.02.008.

Moauro, F. and G. Savio. 2005. "Temporal Disaggregation Using Multivariate Structural Time Series Models." *The Econometrics Journal* 8(2): 214–234. Doi: http://dx.doi.org/10.1111/j.1368-423X.2005.00161.x.

Morley, J.C., C.R. Nelson, and E. Zivot. 2003. "Why Are the Beveridge-Nelson and Unobserved Components Decompositions of GDP so Different?" *Review of Economics and Statistics* 85(2): 235–243. Available at: http://www.jstor.org/stable/3211575 (accessed December 2017).

Ohanian, L.E. and A. Raffo. 2012. "Aggregate Hours Worked in OECD Countries: New Measurement and Implications for Business Cycles." *Journal of Monetary Economics* 59(1): 40–56. Doi: https://doi.org/10.1016/j.jmoneco.2011.11.005.

Pfeffermann, D. 1991. "Estimation and Seasonal Adjustment of Population Means Using Data from Repeated Surveys." *Journal of Business and Economic Statistics* 9(2): 163–175.

Pfeffermann, D. and R. Tiller. 2006. "Small-Area Estimation with State-Space Models Subject to Benchmark Constraints." *Journal of the American Statistical Association* 101(476): 1387–1397. Doi: https://doi.org/10.1198/016214506000000591.

Quenneville, B. and C. Gagné. 2013. "Testing Time Series Data Compatibility for Benchmarking." *International Journal of Forecasting* 29(4): 754–766. Doi: https://doi.org/10.1016/j.ijforecast.2011.10.001.

Rünstler, G. 2004. "Modelling Phase Shifts Among Stochastic Cycles." *Econometrics Journal* 7(1): 232–248. Doi: http://dx.doi.org/10.1111/j.1368-423X.2004.00129.x.

Stock, J.H. and M.W. Watson. 2002. "Macroeconomic Forecasting Using Diffusion Indexes." *Journal of Business and Economic Statistics* 20(2): 147–162. Doi: https://doi.org/10.1198/073500102317351921.

Stock, J.H. and M.W. Watson. 2005. "Implications of Dynamic Factor Models for VAR Analysis." NBER Working Paper 11467, NBER. Doi: http://dx.doi.or/10.3386/w11467.

Stock, J.H. and M.W. Watson. 2011. "Dynamic Factor Models." In *The Oxford Handbook of Economic Forecasting*, edited by M. Clements and D. Hendry, 35–60. Oxford: Oxford University Press.

Tiller, R.B. 1992. "Time Series Modeling of Sample Survey Data from the US Current Population Survey." *Journal of Official Statistics* 8(2): 149–166. Available at: https://www.scb.se/contentassets/ca21efb41fee47d293bbee5bf7be7fb3/time-series-modeling-of-sample-survey-data-from-the-u.s.-current-population-survey.pdf (accessed December 2017).

Valle e Azevedo, J., S.J. Koopman, and A. Rua. 2006. "Tracking the Business Cycle of the Euro Area: a Multivariate Model-Based Bandpass Filter." *Journal of Business and Economic Statistics* 24(3): 278–290. Available at: http://www.jstor.org/stable/27638878 (accessed December 2017).

Wanger, S. 2013. "Arbeitszeit und Arbeitsvolumen in Deutschland – Methodische Grundlagen und Ergebnisse der Arbeitszeitrechnung." *Wirtschafts- und Sozialstatistisches Archiv* 7(1-2): 31–69. Doi: https://doi.org/10.1007/s11943-013-0127-0.

Wanger, S., R. Weigand, and I. Zapf. 2016. "Measuring Hours Worked in Germany – Contents, Data and Methodological Essentials of the IAB Working Time Measurement Concept." *Journal for Labour Market Research* 49(3): 213–238. Doi: https://doi.org/10.1007/s12651-016-0206-0.

Wood, J. and D. Elliott. 2007. "Methods Explained: Forecasting." *Economic and Labour Market Review* 1(12): 55–58. Doi: https://doi.org/10.1057/palgrave.elmr.1410188.

Received August 2015
Revised December 2016
Accepted October 2017