Constructing Coincident Economic Indicators for Emerging Economies

Cem Çakmaklı*
Koç University and University of Amsterdam

Sumru Altug†
Koç University and CEPR

September 30, 2014

Abstract

This paper constructs coincident economic indicators (CEI) of real activity that takes into account the data characteristics for emerging economies. The CEI is derived as the common factor at the monthly interval in a dynamic factor framework that also allows for seasonal factors and uses economic series available at mixed frequencies and different sample periods. The approach is to estimate the density of the indicator, rather than only point prediction, using Bayesian semiparametric estimation. The framework that is used allows for such emerging economy characteristics as changing volatilities and extreme observations, and it is applied for the case of Turkey, an emerging economy that has witnessed changes in the policy regime, crises, and structural shifts over the sample period in question.

Keywords: *Coincident indicator, Bayesian estimation, semiparametric approach, Turkey*

JEL Classification:

*Correspondence to: Cem Çakmaklı, Department of Economics, Koç University, e–mail: ccakmakli@ku.edu.tr, C.Cakmakli@uva.nl. An earlier version of this paper was presented at the Conference on Advances in Applied Macro-Finance and Forecasting at Bilgi University in Istanbul, September 2014.

†Koç University and CEPR, e–mail: saltug@ku.edu.tr.

1 Introduction

The comovement in the behavior of a large number of economic series was noted by Burns and Mitchell (1946) in their quest to define a "business cycle". The notion of a coincident indicator to capture such comovement in macroeconomic time series follows from this approach. Coincident and leading indicators provide timely information about the underlying state of the economy and help to forecast it. Following the suggestion in Sargent and Sims (1977) and others, the dynamic latent factor model has been used to derive coincident and leading indicators in terms of a single common factor summarizing the behavior of a large number of macroeconomic series.

A number of papers have derived coincident economic indicators (CEI) for the US economy using data available at different frequencies. Stock and Watson (1989) construct a monthly coincident indicator of real activity using monthly data on series included in the Department of Commerce's coincident index, namely, industrial production, real personal income less transfer payments, real manufacturing and trade sales, and employee hours in nonagricultural establishments. Likewise, Diebold and Rudebusch (1996), Chauvet (1998), Kim and Nelson (1998) have developed monthly coincident indicators using a monthly data set. Mariano and Murasawa (2003), Aruoba *et al.* (2009) develop (weekly) monthly coincident indicators using a mixed frequency dataset. Banbura *et al.* (2013) and other papers implement nowcasting using a mixed frequency data set.[1] Some recent papers have also developed financial conditions indices (FCI) to examine the role of financial factors in determining future real activity. Hatzius *et al.* (2010) provide an extensive review and comparison of the alternative indices that are available, and note the instability of FCI's across

[1]These papers typically extract a common factor describing the comovement in a set of series using the dynamic factor model, and employ alternative probability models to generate the turning points in the underlying measure of real economic activity. Of these papers, Diebold and Rudebusch (1996), Chauvet (1998), and Kim and Nelson (1998) modify the dynamic factor model to allow for a regime-switching factor based on the Markov switching model proposed by Hamilton (1989).

different sample periods.[2]

In this paper, we construct a measure of economic activity, i.e. a coincident economic indicator, using econometric techniques designed to take emerging market characteristics into account. Specifically, the "dynamic factor model" utilizes the assumption that the variation in a large data set constructed by N coincident variables can be summarized using a single common factor, which will serve as a coincident economic indicator (CEI) in our case. The proposed model for constructing the coincident economic indicator can be cast into the state-space framework. The state space representation involves measurement equations linking the coincident variables to a latent factor, and transition equations describing the dynamics of the unobserved factor. The latent factor is assumed to represent the unobserved state of the economy, and to provide information on its evolution. The methods in question are used to estimate a monthly coincident indicator for a key emerging market economy, namely, Turkey.

One of the main obstacles for deriving a coincident indicator for emerging economies has to do with data availability. Put differently, the information set is limited to only a few variables. The series used to construct the coincident economic indicator in our application are similar to those included in the coincident economic indicator developed by the Department of Commerce for the U.S. economy. Real GDP and the Industrial Production Index (IPI) measured at the quarterly and monthly frequencies, respectively, are included as the key indicators on which most estimates of cyclical activity are based. Total nonagricultural employment provide a measure of labor market conditions while real consumption expenditures, a trade and services turnover index, and a retail sales volumes index help to capture demand-side influences. However, such series for emerging economies are available only at low fre-

[2] They construct an FCI that includes a large array of quantitative and survey-based measures, in addition to standard financial variables such as interest rates and asset prices, and show that it is more tightly related to future economic conditions than other available FCI's.

quencies, they are subject to many revisions, or they span only a short time period.[3]. Therefore, one needs techniques sophisticated enough to handle mixed frequencies and missing observations in a statistically consistent way. As Aruoba *et al.* (2009) note, the state-space framework allows a convenient way to incorporate relevant data from all frequencies and with different time spans. In our case, the series that are available represent a patchwork rather than combination of mixed frequencies, which tends to be a characteristic of typical emerging market data.

Another key aspect of our analysis is that we implement a Bayesian semiparametric approach to estimate the entire density of the indicator, rather than only engaging in point prediction. The framework that is used allows for such emerging economy characteristics as changing volatilities and extreme observations. Bayesian methods to estimate dynamic factor models with regime switching factors dates to Kim and Nelson (1998). The data are available as original observations, and seasonality is taken into account in a factor framework. Koop *et al.* (2013) derive an FCI that is useful for predicting future GDP growth and unemployment based on a framework that allows for time-varying parameters and stochastic volatility. They implement a dynamic version of model averaging by developing a procedure that weights alternative FCI's at different points in time, where the different FCI's are differentiated by the financial variables that are included in them. We allow for time variation in the mean and the variance of the indicator in a discrete manner using Bayesian semiparametrics. Moreover, in our data sets, we have many variables with the same content but with different releases and/or different frequencies over different time periods. In those cases, we allow the factor loadings to change discretely to capture such changes in the data structure.

Despite the large number of applications of this approach in the context of the

[3]Specifically, for Turkey, data on real GDP are available quarterly from 1987 but the coverage and base year have been changed since 1998. Another case in point refers to nonagricultural employment, which was initially available at the biannual frequency but the frequency was later changed to the quarterly frequency and finally, to the monthly frequency starting from 2005.

US or the countries of the euro area[4], there are relatively few applications for emerging market economies. Atabek *et al.* (2005) construct a composite leading indicator for the Turkish economy using seven demand, supply and policy indicators over the period 1987-2005. Akkoyun and Gunay (2012) create backcasts and nowcasts of Turkish GDP growth by taking into account mixed frequencies (such as quarterly and monthly series), ragged ends (some indicators are published before others, and missing data (data not available for some variables at the beginning of the sample). They use data on the industrial production, imports and exports as well as survey data available in the Purchasing Managers' Index. In a paper that is most closely related to our analysis, Aruoba and Sarikaya (2013) develop a monthly indicator of real economic activity using multiple indicators such as real GDP, industrial production, imports of intermediate goods, electricity production and employment data at mixed frequencies. They employ a dynamic factor model formulated and estimated at the monthly frequency to extract the common factor as the economic indicator.

Our model differs from these approaches in various ways. First, we conduct a Bayesian inference by allowing for the parameter and state uncertainty simultaneously. This is crucial in the current setting where data are of mixed frequency and therefore, subject to missing observations. Second, we model seasonality explicitly rather than employing ad-hoc mechanical procedures for removing it prior to the analysis. By doing so, we capture the common information in the data for modeling seasonality. Third, we use Bayesian semiparametric techniques that allow for changing data patterns due to extreme observations and structural changes in the economy when modeling the economic activity indicator. Finally, we use a much broader data set involving many conventional variables that start at the end of the 1980's but also include other more recently released variables.

The remainder of this paper is as follows. Section 2 presents the model. Section

[4]See Banbura *et al.* (2013) for a discussion of the different applications.

3 describes the estimation approach while Section 4 presents the empirical results. Concluding remarks are in Section 5.

2 The Model

Consider the dynamic factor model that formulates the joint behaviour of a set of stock and flow variables that will be used to extract coincident index, $X_t^{i,S}$ and $X_{i,F}^t$, respectively. In the absence of a prior seasonal adjustment, the data are assumed to be comprised of two components, namely, a unique factor that captures the comovement in the different series (which is the coincident index), C_t, and two unique factors to capture seasonal components at the monthly and the quarterly frequencies, $S_{1,t}$ and $S_{2,t}$, respectively.

The **monthly** coincident factor is assumed to evolve with $AR(p)$ dynamics as

$$\Phi(L)C_t = \eta_{C,t}, \tag{1}$$

where $\Phi(L)$ is the lag operator $(I - L - \ldots L^p)$. For the distribution of $\eta_{C,t}$, we do not specify a parametric distribution but we estimate this distribution along with other parameters using the infinite Markov mixture model (iMMM). Specifically, consider the finite Markov mixture model (fMMM) which, in the limiting case, becomes iMMM. A fMMM consists of latent states where, conditional on these states, the observations are assume dto follow a Normal distribution. Each state variable can take a finite number of states where the number of the states are determined prior to analysis, $s_t = 1, \ldots, K$. Taking time dependence into account, the transitions between states follow a Markov process of first order, $\pi_{ij} = \Pr(s_t = j | s_{t-1} = i)$. Conditional on the type of state in period t, error terms follow a Normal distribution with parameters θ_{S_t}. In our case, θ_{S_t} consists of the mean and the variance parameter, μ_t and $\sigma^2_{\eta_C,t}$, of the error term.[5].

In a typical analysis, the transition probabilities from state i follow Dirichlet distribution and therefore, for the prior specification we also use a Dirichlet distribution of symmetric form with parameter γ/K assigning equal probability for each of transition. We complete the prior specification by assigning the distribution H for the prior distribution of θ. To obtain an iMMM, let $K \to \infty$ to obtain a Dirichlet process prior. One of the obstacles of this approach in the Markov mixture models content is that we assign a different Dirichlet process prior for transition from each state. Therefore, there is no coupling across transitions out of different states since the transition probabilities are given independent priors. To introduce coupling, we can define a modified version of fMMM and then let $K \to \infty$, as

$$\begin{aligned} \pi_i &\sim \text{Dirichlet}(\alpha\beta) \\ \beta &\sim \text{Dirichlet}(\gamma/K) \end{aligned} \tag{2}$$

where π_i denotes transition probabilities out of state i. In this specification distribution of all transition probabilities share a common parameter, β, which itself follows a Dirichlet distribution. As $K \to \infty$ this hierarchical prior approaches to hierarchical Dirichlet process. This prior together with a standard likelihood yields the iMMM. The advantages of iMMM is threefold. First, the parameter values at t can take any values unlike fMMM; second, iMMM spans all the continuous distributions with probability 1 indicating that it is truly a density estimation; and third, changes in the parameters behave in a discrete manner suitable for the low frequency macro-finance data.

For a more technical exposition of the iMMM, let G_0 be a probability distribution function over some parameter space Θ such that $\theta \in \Theta$ and let α be a positive scalar. A distribution G is distributed with a Dirichlet process $DP(\alpha, G_0)$ with

[5]The intercept in equation (1) is suppressed as the mean of the error term

base measure G_0 and precision parameter α if for any measurable J partition of the parameter space, $\theta_{1:J} = (\theta_1, \theta_2, \ldots, \theta_J)$, the distribution $G = G(\theta_{1:J})$ follows a Dirichlet distribution $Dir\left(\alpha G_0(\theta_1), \alpha G_0(\theta_2), \ldots, \alpha G_0(\theta_J)\right)$. Specifically

$$f(\theta_t|\theta_{-t})|G = G = \frac{\alpha}{\alpha + T - 1} G_0(\Lambda) + \sum_{i=1, i\neq t}^{T} \frac{\delta^j(\theta_i)}{\alpha + T - 1}. \tag{3}$$

Dirichlet processes span the entire distribution of the discrete probability functions with probability one. This can be seen from the stick breaking representation of the DP, see Sethuraman (1994), with the following scheme. Suppose a part of $W_j\%$ of a stick with a unit initial length is being constantly broken j times where $W_j \sim Beta(1, \alpha)$. As $j \to \infty$ the unit length of the stick can be written as $1 = \sum_{j=1}^{\infty} \pi_j$. Since any partition can be the result of this stick breaking process, Dirichlet process spans the entire distribution of the discrete probability functions with probability one. This stick-breaking construction is denoted by $\pi \sim \text{GEM}(\alpha)$. When hierarchical DPs are considered the base measure itself follows $DP(\gamma, H)$ with H global base measure. In this case, it turns out that using the stick-breaking construction the base measures can be expressed as $G_0 = \sum_{j'=1}^{\infty} \beta_{j'} \delta_{\theta_{j'}}$ and $G_i = \sum_{j'=1}^{\infty} \pi_{ij'} \delta_{\theta'_j}$ with $\beta \sim \text{GEM}(\gamma)$, $\pi_k = DP(\alpha, \beta)$.

We model seasonality similarly using a factor structure, where the first seasonal factor with monthly seasonal dynamics is given by

$$S_{1,t} = -\sum_{s=1}^{11} S_{1,t-s} + \eta_{S_{1,t}}, \tag{4}$$

and the second seasonal factor with quarterly seasonal dynamics evolves as

$$S_{2,t} = -S_{2,t-3} - S_{2,t-6} - S_{2,t-9} + \kappa_t \eta_{S_{2,t}} \tag{5}$$

where $\kappa_t \in \{0,1\}$.[6]. The coincident variables denoted X_t^i load on the (single) coincident factor and the seasonal factors as

$$X_t^i = \lambda_1^i C_t + \lambda_2^i S_{1,t} + \lambda_3^i S_{2,t} + \varepsilon_t^i. \tag{6}$$

2.1 Temporal aggregation

Following Aruoba *et al.* (2009), we work with detrended series rather than the growth rates as in Banbura *et al.* (2013). This enables us to capture the seasonal components much precisely as taking the growth rates obscures the inference about seasonal components. We distinguish between stock and flows variables in our analysis. Formally, suppose that we would like to construct a coincident indicator as, C_t at a given frequency, say the monthly frequency. As the X_t^i are not observed directly but instead we observe the relevant series at a different frequency as $\tilde{X}_t^i$, we have to model this.

Suppose that X_t^i is the desired higher frequency (unobserved) measurement of the observed raw data at low frequency $\tilde{X}_t^i$. We work with the logarithm of the original series. The stock variables can be expressed as

$$\tilde{X}_t^{i,S} = \begin{cases} \lambda_1^i C_t + \lambda_2^i S_{1,t} + \lambda_3^i S_{2,t} + \varepsilon_t^i & \text{if observed} \\ \text{NA} & \text{otherwise} \end{cases} \tag{7}$$

while the flow variables are expressed as

$$\tilde{X}_t^{i,F} = \begin{cases} \sum_{s=0}^{D_i-1} (\lambda_1^i C_t + \lambda_2^i S_{1,t} + \lambda_3^i S_{2,t} + \varepsilon_t^i) & \text{if observed} \\ \text{NA} & \text{otherwise} \end{cases} \tag{8}$$

[6]The monthly seasonal factor can also serve as a quarterly seasonal factor as the convolution of the monthly seasonal factor to quarterly frequency produces a quarterly seasonal factor. However, in our case this is not sufficient as addition of the separate quarterly factor increases the model fit. Results are available upon request by the authors.

where $D_i = 2$ for quarterly variables and $D_i = 5$ for semi-annual variables.

2.2 The final model

Using these definitions, the state space model is written in compact form as

$$\begin{aligned}
\tilde{X}_t^{i,S} &= \lambda_1^i C_t + \lambda_2^i S_{1,t} + \lambda_3^i S_{2,t} + \varepsilon_t^i \\
\tilde{X}_t^{i,F} &= \textstyle\sum_{s=0}^{D_i-1} \left(\lambda_1^i C_t + \lambda_2^i S_{1,t} + \lambda_3^i S_{2,t} + \varepsilon_t^i\right), \quad \varepsilon_t^i \sim N(0,\sigma_i^2) \text{ for } i = 1,2, \\
\Phi(L)C_t &= \eta_{C,t} \qquad \eta_{C,t} \sim G(\theta_t) \quad \theta_t|G \sim G(\alpha, G_0) \quad G_0 \sim DP(\gamma, H) \\
S_{1,t} &= -\textstyle\sum_{s=1}^{11} S_{1,t-s} + \eta_{S_{1,t}}, \quad \eta_{S_{1,t}} \sim N(0, \sigma^2_{S_{1,t}}) \\
S_{2,t} &= -S_{2,t-3} - S_{2,t-6} - S_{2,t-9} + \kappa_t \eta_{S_{2,t}}, \quad \eta_{S_{2,t}} \sim N(0, \sigma^2_{S_{2,t}}).
\end{aligned} \tag{9}$$

Thus, C_t is the common factor to capture the coincident indicator, and $S_{1,t}$ and $S_{2,t}$ are the seasonal factors to capture the seasonality embodied in the monthly and quarterly series, respectively.

2.3 Data

To construct the data set, we use four sets of variables along the lines of many applications using US data. These include series on industrial production, employment, trade and sales, and income; see Stock and Watson (1989) or (Kim and Nelson, 1998). Mariano and Murasawa (2003) add Gross Domestic Product to this data set, as it is the common measure of economic activity at the lower frequency. Additionally, since Turkey can be considered as a small open economy, unlike the US, we also add variables related to the trade balance. Accordingly, we construct our data set to capture the covariation in these sets of variables as much as possible. The variables included in our study are Gross Domestic Product at constant prices (rGDP), the industrial production index (IPI), total employment less agricultural employment (ENP), the trade and services turnover index (TST), the retail sales volume index

(RSV), final consumption at constant prices (rFC), the total export quantity index (XQ) and the total import quantity index (MQ).

While these series are subject to different data revisions and base prices (for real variables) in different periods, only IP, XQ and MQ are observed in the entire sample period at the monthly frequency, which starts as early as the January 1986. Despite the fact that earlier observations are likely to be subject to structural changes or breaks, more precise estimates of the CEI and the seasonal factors are likely to be obtained using more information. Hence, it is crucial to use more observations in our application, as we employ unobserved components models with many missing data points. Therefore, we opt to model the breaks and changes in the different series and to use all available information. The detailed information about the data set is provided in Appendix A.

3 Estimation

We use a Bayesian approach for estimation and inference in the dynamic factor model with mixed frequency observations using Markov Chain Monte Carlo (MCMC) techniques. Specifically, we use Gibbs sampling together with data augmentation (see Geman and Geman, 1984; Tanner and Wong, 1987) to obtain posterior results. For estimating the density of the coincident factor, we use the slice sampler of Neal (2003), see also Van Gael *et al.* (2008) and Kalli *et al.* (2011). In Section 3.2 we discuss the specifications of the prior distributions. Finally, in Section 3.3 we outline the Gibbs sampling algorithm for simulating from the posterior distribution.

3.1 The econometric model

The state space model in a compact form is as follows

$$\begin{aligned} X_t &= \mu + HS_t + \varepsilon_t \qquad \varepsilon_t \sim N(0,R) \\ S_t &= FS_{t-1} + \eta_t \qquad \eta_t, \sim N(0,Q), \end{aligned} \tag{10}$$

where X_t includes X_t^i for $i = 1, \cdots, 6$, μ is a vector constants and the state vector S_t includes current and lagged values of the coincident index C_t and the seasonal factors $S_{1,t}$ and $S_{2,t}$. The error vector is composed by the error terms in the stock and flow representations for the different economic indicators included in our study.

3.2 Prior distributions

We start to the specifications of the prior distributions with the specifications related to the iMMM. This includes the specification of the base prior and the specification of the hyperpriors for the precision parameters α and γ. For the base prior H I use a normal-inverse Gamma distribution as follows

$$H(\mu_0, a, \nu, S) = N\!-\!IG(\mu_0, a, \nu, S) \mapsto \mu|\sigma_H^2 \sim N(\mu_0, \frac{\sigma_H^2}{a}) \text{ and } \sigma_H^2 \sim IG(\nu, S), \tag{11}$$

where $E(\mu) = \mu_0$ and $E(\sigma_H^{-2}) = \nu S^{-1}$. Moreover, S can be decomposed further into $S = \nu v$. For the precision parameters of the DPs, α and γ, we use Gamma distributions $\alpha \sim \text{Gamma}(a_\alpha, b_\alpha)$ and $\gamma \sim \text{Gamma}(a_\gamma, b_\gamma)$.

The remaining model parameters do not involve a DP prior. For the prior distribution of the autoregressive parameters in (1), we specify the standard multivariate normal distribution as

$$f(\text{vec}(F)) = N(\mathbf{0}_p, I_p), \tag{12}$$

where vec(F) is the vec operator stacking each columns of F into a single column vector. $\mathbf{0}_9$ is a vector of zeros of dimension p. Unreported results of prior sensitivity analysis suggests that the likelihood information dominates prior information for F and thus, the prior specification is of limited importance for this parameter.

For the factor loadings $\lambda_{i,k}$ for $i = 1, \ldots, N$ and $k = 1, 2, 3$ we use Normal distributions, $N(1, 10)$. The large value of the variance reflects our ignorance about the parameter value, while the mean is centered around one.

For the variance of the measurement errors, we use uninformative Jeffrey's priors as $f(\sigma_i^2) \propto \frac{1}{\sigma_i^2}$.

Finally, as factors and their loadings are both unobserved, we need to set identification restrictions to uniquely identify the factors. Therefore, the variances of the state errors are restricted to be one except the coincident factor, see Bai and Wang (2012) for example. We also set the loading of the coincident factor to IP as one, and the loading of the quarterly seasonal factor to IP as zero. We also set the loading of the monthly seasonal factor to rGDP as zero to complete the identification restrictions.[7].

3.3 Sampling scheme

The MCMC algorithm consists of the following steps:

1. Initialize the parameters by drawing θ_t from the base prior H. At step (m) of the iteration

2. Sample $f_t^{(m)} = (C_{1:T}^{(m)}, S_{1,1:T}^{(m)}, S_{2,1:T}^{(m))'}$ from $p(f_t | y_{1:T}, \theta_{1:T}^{(m)}, \Phi^{(m)})$ using Carter *et al.* (1994) or another alternative.

3. Sample $\Phi^{(m)}$ from $p(\Phi | C_{0:T}^{(m-1)}, \theta_{1:T}^{(m-1)})$.

[7] We also set other identification schemes such as setting unity for the unconditional variance of the factor. The current identification scheme resulted in more intuitive results in line with stylized facts about Turkish economy. Results are available upon request.

4. Sample $\sigma_i^{2,(m)}$ from $p(\sigma_i^2|y_{1:T}, f_t^{(m-1)})$ using the observations that are not missing.

5. Sample $\lambda_i^{(m)} = (\lambda_{i,1}^{(m)}, \lambda_{i,2}^{(m)}, \lambda_{i,3}^{(m)})'$ from $p(\lambda_i|y_{1:T}, f_t^{(m-1)})$ using the observations that are not missing.

6. Sample $\theta_{1:T}^{(m)}$ from $p(\theta_t|C_{1:T}^{(m-1)}, \theta_{-t}^{(m-1)}, \Phi^{(m)})$ for $t = 1, 2, \ldots, T$ using slice sampler of Neal (2003), see also Van Gael *et al.* (2008) and Kalli *et al.* (2011).

7. Sample $\alpha^{(m)}$ and γ from $p(\alpha|\theta_{1:T}^{(m)})$ and $p(\gamma|\theta_{1:T}^{(m)})$.

8. Repeat (2)-(6) M times.

4 Estimation Results

In this section, we present estimates of the coincident economic indicator (CEI) as well as estimates of the seasonal factors and the time-varying intercept and variance of the CEI.

4.1 The cyclical behavior of the CEI

Figures 1 and 2 present the estimated coincident indicator or common factor based on GDP and IP data, respectively, together with the 95% confidence interval and the shaded recession dates obtained from the BBQ algorithm for the Turkish economy, see Altug and Bildirici (2012). We observe that the CEI is able to track the recessions associated with major financial crises such as those in 1994-1995, the global financial crisis in 2008-2009 as well as earlier recessions associated with the first Gulf War in 1991 and Russian sovereign debt default in 1998. The CEI based on the IP index displays very similar behavior. These recession dates also line up with the dates reported by Aruoba and Sarikaya (2013) based on their indicator of economic activity for Turkey.

As Aruoba and Sarikaya (2013) note, there is considerable volatility in the behavior of the CEI over the sample period, with crossings along the zero line occurring at frequent intervals. We also observe that despite the existence of negative growth in the CEI over several months, the only episodes for which these are significantly different from zero (as signified by the upper band of the 95% confidence interval going below zero) for sustained periods are for the 1994-95, 2000-2001 and 2008-2009 recessions.

4.2 Seasonal factors

Figures 3 and 4 display the behavior of the seasonal indicator at monthly and quarterly frequencies while Table 1 shows the factor loadings on the monthly and quarterly factors. The construction of these factors jointly with the CEI is a contribution of our framework, which makes use of unadjusted data in recovering jointly the CEI and the seasonal factors. The monthly factor is more variable and there is a tendency for this variability to increase in the 2000's. By contrast, the quarterly seasonal factor shows the opposite behavior, with larger fluctuations during the 1990's. These results indicate that *ad hoc* seasonal adjustment methods may obscure important features of the time variation in the underlying series, and hence, lead to mis-specification bias in estimated unobservable factors intended to capture the common movement in key aggregate series. Finnally, from Table 1, we observe that the monthly series such as XQ and MQ load more on monthly seasonal factor denoted by λ_{S_1} while the quarterly series such as rFC, THC load more on quarterly seasonal factor denoted by λ_{S_2}.

4.3 Factor loadings

Figures 5-7 display the factor loadings for the non-agricultural employment and real GDP series. The first series is constructed from data at the bi-annual, quarterly and monthly frequencies while the real GDP experienced a change in the base year in 1998. We observe significant differences in the factor loadings on the common factor for both series. For the non-agricultural employment series, we observe that the factor loading on the common factor is not significantly different from zero using series measured at the bi-annual frequency up to 2000:B2. After this date, the series loads with a significantly negative coefficient on the common factor up until 2004:Q4, and a significantly positive coefficient for the remainder of the sample period.

For the real GDP series, the factor loading on the common factor barely fluctuates around a value of 0.20 for the entire sample period. However, the confidence band around this value os much larger for pre-1998 period. Finally, we observe that the data revisions in the real GDP series are associated with different factor loadings on the quarterly seasonal component, further justifying our approach of multiple factor loadings to account for different data releases over the sample period.[8]

4.4 Time-varying intercepts and volatility

The estimation algorithm allows for time variation in the intercept and the variance of the unobserved CEI denoted by C_t. Figure 8 shows that are minor changes in the intercept and increasing volatility during recessions. In particular, we observe that the intercept falls and volatility increases during the 1994-1995, 2000-2001 and 2008-2009 crises for the Turkish economy. There is also a spike in the volatility series during the second half of 2013, when Turkey experienced mass protests known as the Gezi protests as well as allegations of corruption against senior political figures.

[8]While the real consumption expenditures series rFC also experiences a new release in 1998, we did not report the estimated factor loadings across the different periods because they were nearly identical and equal to the values reported in Table 1. The results are available upon request.

5 Conclusion

The challenges of constructing a coincident indicator or real activity for emerging economies such as Turkey stem from different sources. One of these challenges has to do with the sizable seasonal components in the underlying series, which may include a common component. The second challenge has to do with the data availability and the mismatch of frequencies and sample periods for the existing series. A third challenge is associated with the changing nature of trends and volatility in the behavior of an underlying measure of real activity, which may reflect changes in policy regimes, the role of external shocks, and structural adjustment in the domestic economy itself.

The main findings of the paper can be summarized as follows. The coincident indicator which is constructed from the common unobservable common factor is able to capture the timing, duration and severity of recessions for the Turkish economy. However, we observe significant variation in the factor loadings on the common factor for series in which there are changes in the data structure throughout the sample period. Such variation provides justification for our approach of accounting for changes in the nature of the series due to data releases, a feature that is typically absent from other similar studies. The monthly and quarterly common seasonal factors are sufficient for capturing seasonal patterns in the series while semiparametric structure allows for sharper turning points.

In future work, we seek to extend the coincident indicator at a higher frequency and conduct density nowcasting using semiparametric structure. We also seek to include financial variables such as the slope of the yield curve or changes in the supply of credit to capture the role of financial variables on future real economic activity. Nevertheless, our approach provides a potential template for developing a coincident economic indicator for emerging market economies in that it allows for mixed frequencies, missing data and time-varying intercepts and volatility in the underlying common factor. Such features address the problems of data availability for such economies as well as accounting for potential regime shifts, structural breaks and geo-political risks.

References

Akkoyun, C. and M. Gunay (2012), Nowcasting Turkish GDP Growth, Working Paper Series 12, Central Bank of the Republic of Turkey.

Altug, S. and M. Bildirici (2012), Business Cycles in Developed and Emerging Economies: Evidence from a Univariate Markov Switching Approach, *Emerging Markets Finance and Trade, M.E. Sharpe, Inc.*, 48, 4–38.

Aruoba, B., F. Diebold, and C. Scotti (2009), Real-Time Measurement of Business Conditions, *Journal of Business & Economic Statistics*, 27, 417–427.

Aruoba, B. and C. Sarikaya (2013), A Real Economic Activity Indicator for Turkey, *Central Bank Review*, 13.

Atabek, A., E. Cosar, and S. Sahinoz (2005), A New Composite Leading Indicator for Turkish Economic Activity, *Emerging Markets Finance and Trade*, 41, 45–64.

Bai, J. and P. Wang (2012), Identification and estimation of dynamic factor models, MPRA Paper 38434, University Library of Munich, Germany.

Banbura, M., D. Giannone, M. Modugno, and L. Reichlin (2013), Now-Casting and the Real-Time Data-Flow, in G. Elliott and A. Timmermann (eds.), *Handbook of Economic Forecasting*, Elsevier-North Holland.

Burns, A. and W. Mitchell (1946), *Measuring Business Cycles*, NBER.

Carter, C., , and R. Kohn (1994), On Gibbs Sampling for State Space Models, *Biometrika*, 81, 541–553.

Chauvet, M. (1998), An Econometric Characterization of Business Cycles with Factor Structure and Markov Switching, *International Economic Review*, 39, 969–96.

Diebold, F. and G. Rudebusch (1996), Measuring Business Cycles: A Modern Perspective, *Review of Economic and Statistics*, 78, 67–77.

Geman, S. and D. Geman (1984), Stochastic Relaxations, Gibbs Distributions, and the Bayesian Restoration of Images, *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 6, 721–741.

Hamilton, J. D. (1989), A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle, *Econometrica, Econometric Society*, 57, 357–84.

Hatzius, J., P. Hooper, F. Mishkin, K. Schoenholtz, and W. M. (2010), Financial Conditions Indexes: A Fresh Look after the Financial Crisis, Working paper.

Kalli, M., J. E. Griffin, and S. G. Walker (2011), Slice sampling mixture models, *Statistics and computing*, 21, 93–105.

Kim, C. and C. Nelson (1998), Business Cycle Turning Points, A New Coincident Index, And Tests Of Duration Dependence Based On A Dynamic Factor Model With Regime Switching, *Review of Economic and Statistics*, 80, 188–201.

Koop, G., P. Hooper, and D. Korobilis (2013), A new Index of Financial Conditions, Working paper.

Mariano, R. S. and Y. Murasawa (2003), A New Coincident Index of Business Cycles Based on Monthly and Quarterly Series, *Journal of Applied Econometrics*, 18, 427–443.

Neal, R. (2003), Slice Sampling, *Annals of Statistics*, 31, 705–767.

Sargent, T. and C. Sims (1977), Business Cycle Modeling Without Pretending to Have Too Much A Priori Economic Theory, in *New Methods in Business Cycle*

Research: Proceedings from a Conference, Federal Reserve Bank of Minneapolis, pp. 45–109.

Sethuraman, J. (1994), A constructive definition of Dirichlet priors, *Statistica Sinica*, 4, 639–650.

Stock, J. and M. Watson (1989), New Indexes of Coincident and Leading Economic Indicators, in *NBER Macroeconomics Annual 1989, Volume 4*, National Bureau of Economic Research, Inc.

Tanner, M. A. and W. H. Wong (1987), The Calculation of Posterior Distributions by Data Augmentation, *Journal of the American Statistical Association*, 82, 528–550.

Van Gael, J., Y. Saatci, Y. W. Teh, and Z. Ghahramani (2008), Beam Sampling for the Infinite Hidden Markov Model, in *Proceedings of the 25th International Conference on Machine Learning*, ICML '08, ACM, New York, NY, USA.

Appendix A The data set

We use seasonally unadjusted series with different release combinations, where the detailed information is provided in the following Table ??.

Variable name	Mnemonic	Frequency	Release
Industrial production index	IP	Monthly	1986M1:2014M5
Import quantity index	MQ	Monthly	1986M1:2014M5
Export quantity index	XQ	Monthly	1986M1:2014M5
Gross Domestic Product	rGDP	Quarterly	1987Q1:1997Q1 1998Q1:2014Q1
Final Consumption	rFC	Quarterly	1987Q1:1997Q1 1998Q1:2014Q1
Total employment less agricultural employment	ENP	Biannual quarterly monthly	1987B1:1999B2 2000Q1:2004Q4 2005M1:2014M5
Trade and services turnover index	TST	Quarterly	2005Q3:2014Q1
Retail sales volume index	RSV	Monthly	2010M1:2014M5

Table A.1: Data on economic indicators: Frequency of observation and sample periods

	λ_{S_1}			λ_{S_2}		
	5% LB	Median	5% UB	5% LB	Median	5% UB
IP	0.74	0.88	1.02	0	0	0
rFC	-0.12	-0.06	0.00	0.82	0.93	1.07
XQ	0.73	0.99	1.26	-0.58	-0.32	-0.08
MQ	0.87	1.10	1.34	-0.29	-0.08	0.13
RT	1.32	1.32	1.32	0.64	0.64	0.65
THC	0.22	0.76	1.26	0.90	1.17	1.49

Table 1: Factor loadings on the monthly and quarterly seasonal factors

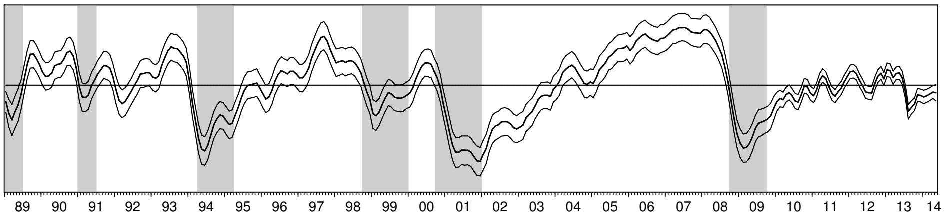

Figure 1: Coincident factor and GDP based recession dates (BBQ algorithm)

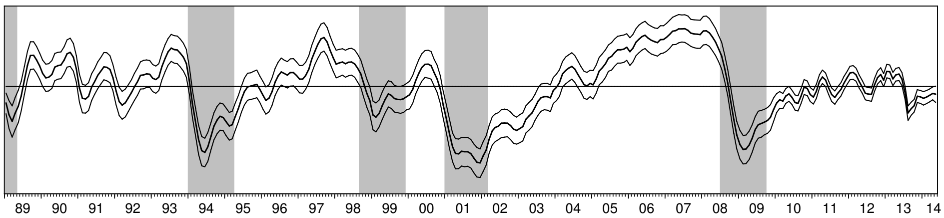

Figure 2: Coincident factor and IP based recession dates (BBQ algorithm)

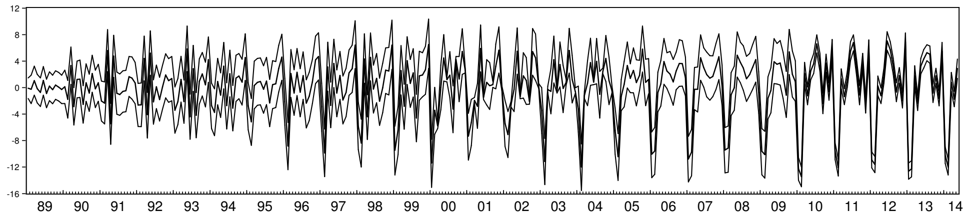

Figure 3: Monthly factor

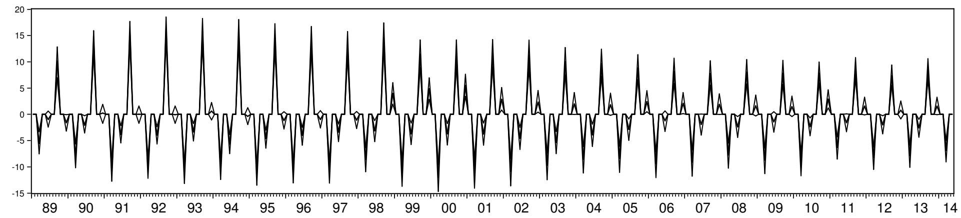

Figure 4: Quarterly factor

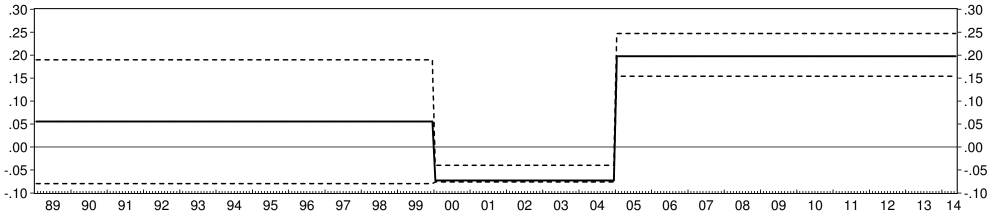

Figure 5: Factor loadings of non-agriculture employment series

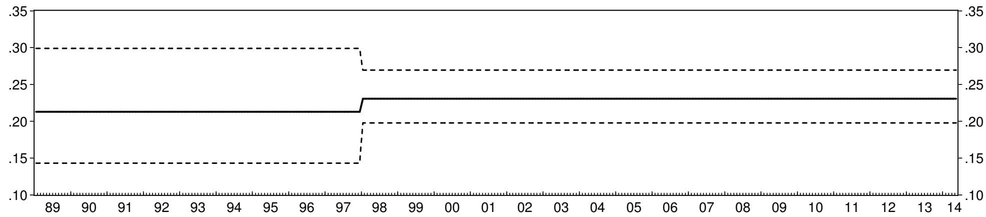

Figure 6: Factor loadings of GDP series

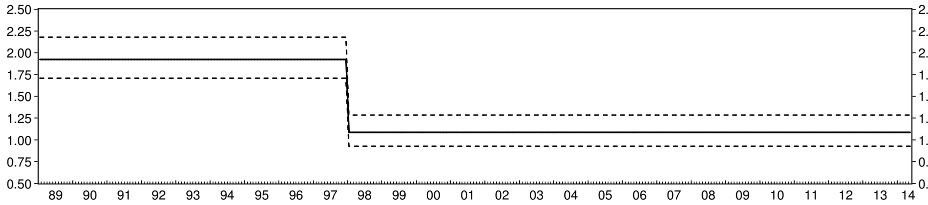

Figure 7: Factor loadings of GDP series for quarterly seasonal factor

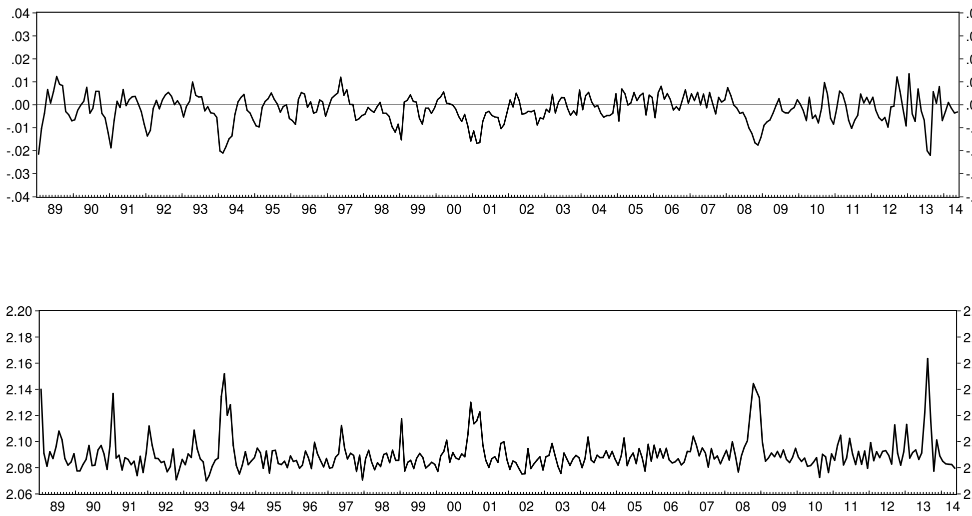

Figure 8: Time-varying intercept and volatility of the coincident indicator