L'Actualité économique

Selection of the Number of Factors in Presence of Structural Instability: A Monte Carlo study

Charles Olivier Mao Takongmo et Dalibor Stevanovic

Volume 91, numéro 1-2, mars–juin 2015

Identification, Simulation and Finite-Sample Inference

URI : https://id.erudit.org/iderudit/1036919ar
DOI : https://doi.org/10.7202/1036919ar

Aller au sommaire du numéro

Éditeur(s)

HEC Montréal

ISSN

0001-771X (imprimé)
1710-3991 (numérique)

Découvrir la revue

Citer cet article

Mao Takongmo, C. O. & Stevanovic, D. (2015). Selection of the Number of Factors in Presence of Structural Instability: A Monte Carlo study. *L'Actualité économique*, *91*(1-2), 177–233. https://doi.org/10.7202/1036919ar

Résumé de l'article

In this paper we study the selection of the number of primitive shocks in exact and approximate factor models in the presence of structural instability. The empirical analysis shows that the estimated number of factors varies substantially across several selection methods and over the last 30 years in standard large macroeconomic and financial panels. Using Monte Carlo simulations, we suggest that the structural instability, expressed as time-varying factor loadings, can alter the estimation of the number of factors and therefore provides an explanation for the empirical findings.

L'Actualité économique, Revue d'analyse économique, vol. 91, n[os] 1-2, mars-juin 2015

SELECTION OF THE NUMBER OF FACTORS IN PRESENCE OF STRUCTURAL INSTABILITY: A MONTE CARLO STUDY*

Charles Olivier MAO TAKONGMO
Department of Economics
McGill University
maotcharles@yahoo.fr

Dalibor STEVANOVIC
Département des sciences économiques
Université du Québec à Montréal
dstevanovic.econ@gmail.com

Abstract–In this paper we study the selection of the number of primitive shocks in exact and approximate factor models in the presence of structural instability. The empirical analysis shows that the estimated number of factors varies substantially across several selection methods and over the last 30 years in standard large macroeconomic and financial panels. Using Monte Carlo simulations, we suggest that the structural instability, expressed as time-varying factor loadings, can alter the estimation of the number of factors and therefore provides an explanation for the empirical findings.

Introduction

Following the improvement of information technology, large panels of economic and financial time series are now available. Using a large data set in econometric analysis can lead to the curse-of-dimensionality problem. One such example is the rise in degrees-of-freedom when the number of variable increases. On the other hand, choosing among variables introduces an element of arbitrariness and can lead to misspecification and misleading results (see Hansen and Richard, 1987; Ludvigson and Ng, 2007). A solution to this problem is to use the factor analysis where the information in hundreds of economic and financial time series can be

* We thank an anonymous referee and the Editor Marie-Claude Beaulieu for useful discussions and comments. The second author acknowledges financial support from the *Fonds de recherche sur la société et la culture* (Québec)

summarized by a relatively small number of (common and latent) factors (see, among others, Chamberlain and Rothschild, 1983; Connor and Korajczyk, 1986, 1988, 1993; Chen, Roll and Ross, 1986; Stock and Watson, 2002; Bernanke, Boivin and Eliasz, 2005; Ludvigson and Ng, 2007, 2009, 2011).

The existence of the factor model is strictly related to the number of primitive shocks in a data set (Rao, 1955). The choice of the number of factors is very important. In fact, researchers can face misspecification problems when the number of factors is underestimated, or problems related to power when the number of factors is overestimated. Many methods have been proposed to estimate the number of (static and dynamic) factors (see, among others, Bai and Ng, 2002, 2007; Onatski, 2009, 2010; Alessi, Barigozzi and Capasso, 2010; Ahn and Horenstein, 2013; Hallin and Liska, 2007; Amengual and Watson, 2007).

The aim of this paper is to study the performance, in terms of the selection of the number of factors, of different tests and information criteria in the context of structural instability. First, we conduct an extensive comparison of all the procedures using several large macroeconomic and financial panels. The empirical results shows that: i) the estimated number of factors differs substantially across the selection methods; ii) it varies a lot over time across, and within, selection methods. Several explanations are possible. The factors (often perceived as states of economy) become more or less pervasive over time such that their dimension can be harder to estimate. The structural changes, such as adoption of new monetary and fiscal policies, can affect the way observable series load on the factors.

In the second part of the paper, we perform many Monte Carlo simulations to suggest that the structural instability can alter the estimation of the number of factors and therefore explain the empirical findings above. In particular, we approximate the structural changes by allowing for time-varying factor loadings. Our work is related to Bates, Plagborg-Moller, Stock and Watson (2013), BPSW hereafter, who consider the estimation of the factor space in the presence of time variation in factor loadings. They also verify the performance of the Bai and Ng (2002) criterion to successfully predict the dimension of the factor space. The second related paper is Chen, Dolado, and Gonzalo (2014), who provide a framework to test for large breaks in factor loadings[1]. They also show that the Bai and Ng (2002) information criteria are likely to overestimate the true number of factors in the presence of large breaks. Finally, Guo-Fitoussi and Darné (2014) concentrate on comparing finite sample properties among many selection rules. Our contribution to this literature consists of: i) providing empirical evidence for the time varying factor structure, in terms of the number of factors, in macroeconomic and large financial data sets; ii) assessing the performance of several selection rules in the presence of irregularities discussed above. In addition, we study the robustness of selection methods in small and large samples, and in exact and approximate factor structures.

1. See also Breitung and Eickmeier (2011) who test for the presence of structural breaks in dynamic factor models.

The results from our extensive simulation exercise show that structural instabilities, taking several forms of time-variant factor loadings, together with cross-sectional and time dependence of the idiosyncratic component, do alter the estimation of the number of factors across all six most popular selection methods used in the literature. These results can provide an explanation to the empirical evidence on large volatility in the estimated number of factors.

In the first section, we present the time-varying parameters factor model framework. The selection rules considered in our analysis are shown in Section 2. The empirical part of the paper is presented in Section 3. The Monte Carlo simulation experiments are detailed in Section 4. Additional empirical results are presented in the Appendix.

1. The Factor Model

In this framework, the large number of observed time series are modelled as dependent on a small number of latent factors. The factor model can be written as follows:

$$X_{i,t} = \lambda'_{i,t} F_t + e_{i,t}, \quad i = 1,\dots,N \quad t = 1,\dots,T \tag{1}$$

where $X_{i,t}$ is the observed data, $\lambda_{i,t} \in \mathbb{R}^q$ is the possibly time varying factor loading, F_t is a q x 1 vector of latent common factors and $e_{i,t}$ is an idiosyncratic error assumed to be uncorrelated with F_t at all leads and lags.

Define $\mathbf{X_t} = (\mathrm{X_{1,t}},\dots,\mathrm{X}_{N,t})'$, $\wedge_\mathbf{t} = (\lambda_{\mathbf{1,t}},\dots,\lambda_{N,t})'$, $\mathbf{e_t} = (\mathrm{e_{1,t}},\dots,\mathrm{e}_{N,t})'$, and $\mathbf{F_t} = (\mathrm{F_{1,t}},\dots,\mathrm{F}_{q,t})'$ such that the model can be written in a more compact form:

$$\mathbf{X_t} = \wedge_\mathbf{t} \mathbf{F_t} + \mathbf{e_t}. \tag{2}$$

Following BPSW the structural instability may be introduced by modelling the factor loadings as follows:

$$\wedge_t = \wedge_0 + h_{NT} \zeta_t, \tag{3}$$

where h_{NT} is a deterministic scalar sequence that may depend on N and T. h_{NT} sets the scale of deviation. ζ_t is a possibly random process of dimension N x r. ζ_t will be modelled depending on which type of instability we want to assess. For example, ζ_t may be white noise, in which case the factor loading $\wedge_t$ will be the initial loading matrix $\wedge_0$ plus uncorrelated noise. ζ_t may also be modelled as a random walk, which gives a standard continuous time-varying parameter model. Finally, ζ_t may be a single deterministic break. Of course, if $\wedge_t$ is constant, (1) becomes standard factor model with constant parameters.

Note that we only consider the time instability in factor loadings and do not specify a time-varying VAR process for factors, unlike in Korobilis (2013) and Eickmeier, Lemke and Marcellino (2014). Our goal is not to study how impulse responses of X_t are changing over time, but to verify if the estimation of the number of factors is affected by structural instabilities in the way the observable series are linked to latent states of the economy.

2. Tests and Criteria for Selecting the Number of Factors

We consider several selection methods that have been recently developed in approximate static linear factor model framework. In this section, they are presented briefly, the details can be found in the original references. Information criteria procedures are represented by Bai and Ng (2002) and Amengual and Watson (2007). Onatski (2010) and Ahn and Horenstein (2013) are tests based on the theory of random matrices, while Bai and Ng (2007) exploit the rank of matrices. Finally, Hallin and Liska (2007) build on spectral density representation of factor models. Some of these procedures are suited for selecting the number of static factors and others seek to determine the number of dynamic factors. In our simulation designs, we only consider the case where the number of static and dynamic factors is the same, i.e. the two representations are equivalent.

2.1 *Bai and Ng (2002)*

Information criteria select the number of factors which minimizes the variance explained by the idiosyncratic component. The estimated number of factor is:

$$\hat{k} = \underset{0 \le k \le r_{max}}{argmin}\left(\left[\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\left(X_{i,t} - \hat{\lambda}_i^{k'}\hat{F}_t^k\right)^2\right] + kp(N,T)\right), \tag{4}$$

where $\hat{\lambda}_i^k$ and $\hat{F}_t^k$ are the principal components analysis estimators of the factor loadings and factors, when the number of static factors is k. $p(N,T)$ is a penalty function that is used to avoid over-parametrization. The authors provide 16 different specifications of the objective function. The most popular one that we consider in the rest of the paper is the IC_{p2} .

2.2 *Amengual and Watson (2007)*

Assume, in addition to the observational equation (1), that F_t follow a finite VAR process:

$$F_t = \sum_{i=1}^{p}\Phi_i F_{t-i} + \varepsilon_t. \tag{5}$$

Let η_t represents the vector of q common *dynamic* shocks. The innovation ε_t can be written as $\varepsilon_t = G\eta_t$, where G is k x q with full column rank. By substitution, we have:

$$\mathbf{e_{Xt}} = \mathbf{X_t} - \sum_{\mathbf{i=1}}^{\mathbf{p}} \wedge \Phi_\mathbf{i}\mathbf{F_{t-i}} = \wedge\mathbf{G}\eta_\mathbf{t} + \mathbf{e_t}\,. \tag{6}$$

Hence, e_{Xt} follows a static factor model with q factors that correspond to the common shocks η_t. In practice, e_{xt} is obtained by the following calculations:

$$\hat{e}_{Xt}^{A} = X_t - \sum_{i=1}^{p} \hat{\wedge} \hat{\Phi}_i \hat{F}_{t-i}, \tag{7}$$

$$\hat{e}_{Xt}^{B} = X_t - \sum_{i=1}^{p} \hat{\Pi}_i^{ols} \hat{F}_{t-i}, \tag{8}$$

where $\hat{F}$ and $\hat{\wedge}$ denote the principal components estimators of F and $\wedge$, using the consistent estimator of k, and $\left(\hat{\Phi}_1 \hat{;} \Phi_2, \ldots \hat{;} \Phi_p\right)$ the ordinary least square estimator of $\hat{F}_t$ onto $\left(\hat{F}_{t-1}, \hat{F}_{t-2}, \ldots, \hat{F}_{t-p}\right)$. On the other hand, $\left(\hat{\Pi}_1^{ols} \hat{;} \Pi_2^{ols}, \ldots \hat{;} \Pi_p^{ols}\right)$ are the OLS estimators from projection of X_t onto $\left(\hat{F}_{t-1}, \hat{F}_{t-2}, \ldots, \hat{F}_{t-p}\right)$.

Finally, the Bai and Ng (2002) IC_{p2} criteria are applied on an estimate of e_{X_t} to select the number of dynamic common shocks. In our exercises, we concentrate only on static factor models so the matrix G is identity, and we will use the estimator in (7).

2.3 *Onatski (2010)*

Onatski (2010) develops an estimator of the number of factors–in the approximated factor models–that performs well even when the idiosyncratic terms are correlated. Assume that the idiosyncratic components of the data can been written as $e = A\varepsilon B$, where A and B are two largely unrestricted matrices and ε is an N x T matrix with i.i.d. Gaussian errors. Both (limited) cross-sectional and temporal correlations in e are allowed. Onatski (2010) observes that any finite number of the largest idiosyncratic eigenvalues of the sample covariance matrix clusters around a single point, while all the systematic eigenvalues–the number of which equals the number of factor–diverge to infinity. The estimator then separates the diverging eigenvalues from the cluster and counts the number of separated eigenvalues–this is the estimated number of factors. Bai and Ng (2002), Hallin and Liska (2007), and Amengual and Watson (2007) made the assumption that the factor's cumulative effect on the N cross-sectional units grows proportionally to N. According to this assumption, with r static factors, r eigenvalues of the data's covariance matrix grow proportionally to N while the rest of the eigenvalues stay bounded. Onatski (2010) estimates the number of factors without making any assumption on the rate of growth of the factor's cumulative effect.

Let k be the number of factors, and λ_j the j largest eigenvalues of XX'/T, Onatski (2010) shows that for j > k, the differences $\lambda_j - \lambda_{j+1}$ converge to zero while the differences $\lambda_k - \lambda_{k+1}$ diverge to infinity. Let $\left\{k_{max}^n, n \in \mathbb{N}\right\}$ be a slowly increasing sequence of real numbers such that $\left(k_{max}^n / n\right) \to 0$ as $n \to \infty$. The family of estimators is defined as:

$$\hat{k}(\delta) = max\left\{i \leq k_{max}^n : \lambda_i - \lambda_{i+1} \geq \delta\right\},$$

where k_{max}^n is the maximum possible number of factors having a sample of size n.

2.4 *Ahn and Horenstein (2013)*

The idea is based on the fact that the k largest eigenvalues of the variance matrix of N response variables grow unboundedly as N increases, while the other eigenvalues remain bounded. The estimators are obtained by maximizing the ratio of two adjacent eigenvalues. The two estimators are:

$$\hat{k}_{ER} = \underset{0 \le k \le k_{max}}{argmax} \frac{\tilde{\mu}_{NT,k}}{\tilde{\mu}_{NT,k+1}}$$

and

$$\hat{k}_{GR} = \underset{0 \le k \le k_{max}}{argmax} \frac{\log\left[V(k-1)/V(k)\right]}{\log\left[V(k)/V(k+1)\right]},$$

where $V(k) = \sum_{j=k+1}^{m} \tilde{\mu}_{NT,j}$ and $\tilde{\mu}_{NT,k} := \psi_k\left[XX'/(NT)\right]$ are the k^{th} largest eigenvalues of the positive semi definite matrix $XX'/(NT)$. ER refers to the eigenvalue ratio and GR to the growth ratio.

2.5 *Bai and Ng (2007)*

Bai and Ng (2007) exploit the fact that if a r x r matrix Σ_u has rank q, the $k-q$ smallest eigenvalues are zero. Let $c_1 > c_2 > ... > c_N$ be the ordered eigenvalues of Σ_u, and

$$D_{1k} = \left(\frac{c_{k+1}^2}{\sum_{j=1}^{r} c_j^2} \right)^{1/2} \text{ and } D_{2k} = \left(\frac{\Sigma_{j=k+1}^{r} c_{k+1}^2}{\sum_{j=1}^{r} c_j^2} \right)^{1/2}.$$

When the true eigenvalues $c_{q+1},...c_r$ are zero, D_{1k} and D_{2k} should be zero for any $k > q$. The covariance matrix Σ_u is estimated by $\hat{\Sigma}_u = \frac{1}{T-p} \sum_{t=1}^{T} \hat{u}_t \hat{u}_t'$, where $\hat{u}_t$ are the residuals from estimation of the VAR(p) process in $\hat{F}$. The cut-off point is used to account for estimation error.

2.6 *Hallin and Liska (2007)*

Let $\sum_n(\theta)$, $\theta \in [-\pi, \pi]$ represent the spectral density matrices and $\lambda_{n1}(\theta),...\lambda_{nn}(\theta)$ its eigenvalues in decreasing order of magnitude. If the spectral density matrices $\sum_n(\theta)$ are known, Hallin and Liska (2007) propose selecting the number of factors as:

$$\hat{q}_n = \underset{0 \le k \le q_{max}}{argmin} \left[\frac{1}{n} \sum_{j=k+1}^{n} \int_{-\pi}^{\pi} \lambda_{n,j}(\theta) d\theta + kp(n) \right],$$

where $p(n)$ is a penalty function, and q_{max} is some predetermined upper bound. In this case, $\hat{q}_n$ is deterministic because the spectral density matrices $\sum_n(\theta)$ are assumed known. Under assumptions in their paper, if the penalty is such that $\lim_{n \to \infty} p(n) = 0$ and $\lim_{n \to \infty} np(n) = \infty$, we have that $\lim_{n \to \infty} \hat{q}_n = q$.

If the spectral density matrices $\sum_n(\theta)$ are unknown, they can be estimated by the lag window estimator $\sum_n^T(\theta)$:

$$\sum_n^T(\theta) := \frac{1}{2\pi} \sum_{u=-M_T}^{M_T} w(M_T^{-1}u)\Gamma_{n,u}^T e^{-iu\theta},$$

where $x \to w(x)$ is a positive even-weight function and $M_T > 0$ is a truncation parameter, $\Gamma_{n,u}^T$ is the sample cross-covariance matrix of $X_{n,t}$ and $X_{n,t-u}$ based on T information.

The estimated factor number, for a given pair n and T, are:

$$\hat{q}_{1,n}^T = \underset{0 \le k \le q_{max}}{argmin}\left[\frac{1}{n}\sum_{i=k+1}^{n}\frac{1}{2M_T+1}\sum_{l=-M_T}^{M_T}\lambda_{ni}^T(\theta_l) + kp(n,T)\right]$$

or

$$\hat{q}_{2,n}^T = \underset{0 \le k \le q_{max}}{argmin}\left[\log\left(\frac{1}{n}\sum_{i=k+1}^{n}\frac{1}{2M_T+1}\sum_{l=-M_T}^{M_T}\lambda_{ni}^T(\theta_l)\right) + kp(n,T)\right],$$

where $p(n,T)$ is a penalty function, $\theta_l := \pi l / (M_T + 1/2)$ for $l = -M_T,...,M_T$, q_{max} is the predetermined upper bound and the eigenvalues $\lambda_{ni}^T(\theta_l)$ are those of the lag window estimator $\sum_n^T(\theta)$. Under the assumptions in Hallin and Liska (2007), the estimators $\hat{q}_{1,n}^T$ and $\hat{q}_{2,n}^T$ are consistent.

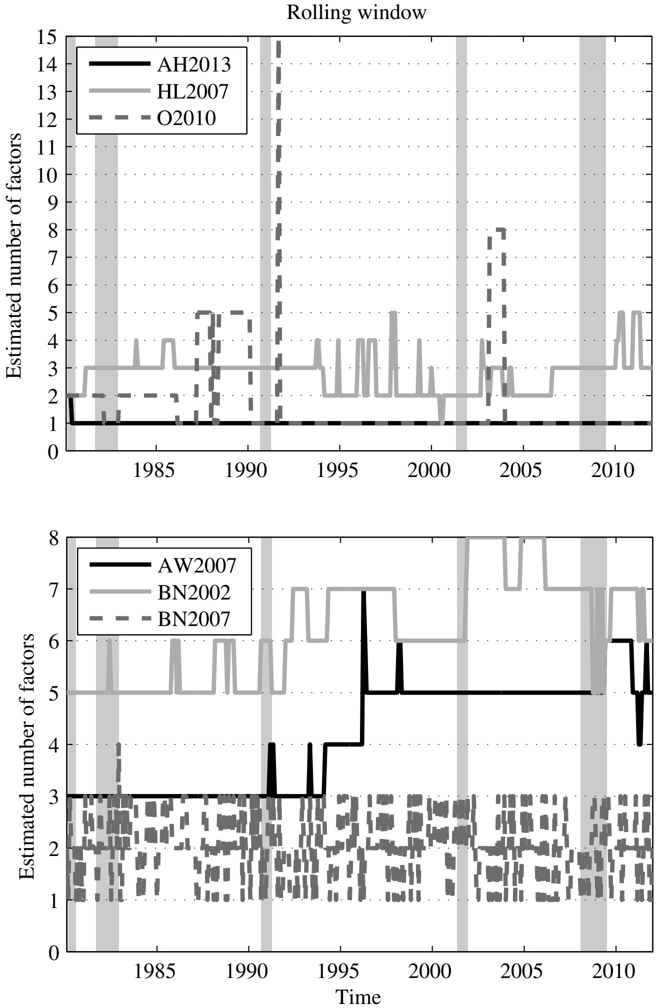
3. Empirical Evidence

Despite a good performance of all selection methods in simulation experiments under regular conditions, their application to large macroeconomic and financial data sets produces mitigated results. In particular, the estimated number of factors varies significantly across the selection procedures and even within a single one.

In this section we compare all the procedures using a variety of macroeconomic and financial panels. A number of conditions can affect the performance of selection methods. First, the macroeconomic panel must be constructed in a way that is representative of economy: time series for different sectors of economic real activities, prices, monetary and credit aggregates, interest rates, etc. The sectoral and disaggregate data are more and more readily available, but adding many series of the same type is not always recommended because it may alter the estimation of common factors, as pointed by Boivin and Ng (2006). The most used US macroeconomic panel is the one from Stock and Watson (2002). While it has been updated by a number of researchers, the core of the data set–in terms of the relative importance of sectors–is always the same. Second, all of these time series must be stationary. In some cases the solution is easy, but in others the transformation

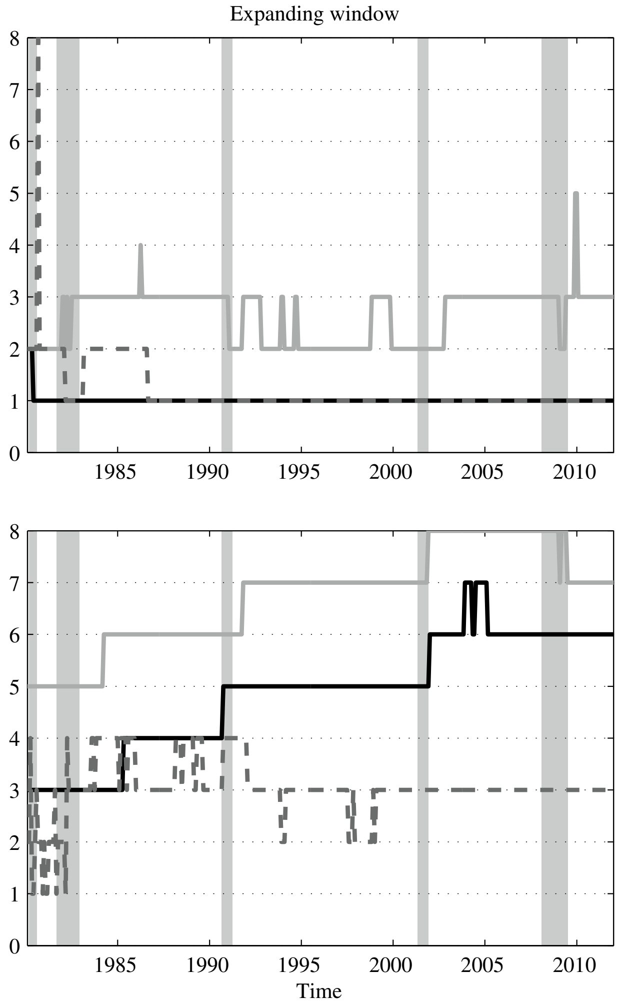
FIGURE 1

NUMBER OF FACTORS OVER TIME: MACROECONOMIC PANEL



to be applied is not obvious. For example, some researchers kept interest rates and inflation rates in levels (Bernanke, Boivin and Eliasz, 2005), while others considered the first difference (Stock and Watson, 2005). Since these series usually represent an important part of the sample, the stationary transformations may substantially modify the correlation structure and hence alter the estimation of the number of factors. Finally, the frequency in which the time series are observed and transformed can be important. Financial indicators are often available on a daily basis while real economic activity series are observed at best monthly. If, in addition, one requires

FIGURE 1 (CONTINUED)



NOTE: This figure presents the selection of the number of factors during 1980-2012 period. The first column presents results computed for a rolling window with 192 months in size (the initial period is 1964M01 - 1979M12). The second column presents results for the expanding window where the time series size grows every period. AH2013 stands for Ahn and Horenstein (2013), HL2007 for Hallin and Liska (2007), O2010 for Onatski (2010), AW2007 for Amengual and Watson (2007), BN2002 and BN2007 for Bai and Ng (2002,2007), respectively. Shaded areas represent the NBER recession periods.

quarterly series such GDP and government spending indicators, the construction of the data set involves several temporal aggregations that are known to change the time series properties (see Lutkepohl, 1984). To investigate the empirical stability of results, we estimate the number of factors in several data sets and across time.

3.1 *Number of Factors in a Large Panel of Macroeconomic Variables*

Figure 1 presents the selection of the number of factors in a large US macroeconomic panel used in Jurado, Ludvigson, and Ng (2013). The data description is available in Appendix B. Essentially, it is an updated version of the Stock-Watson data set, which consists of 132 monthly macroeconomic series observed between 1964M01–2011M12. The data have been stationarized following Stock and Watson (2005): interest, unemployment rates, and inflation measures are in first-difference. We start selecting a number of factors within the 1964M01–1979M12 sub-period and then continue until the end with rolling and expanding windows (first and second column panels, respectively). The first row panels present results for Ahn and Horenstein (2013), Hallin and Liska (2007), and Onatski (2010) procedures while the results for information criteria are presented in the second row. In the case of Bai and Ng (2002), we show the IC_{p2} criterion, which is also used in the second step of Amengual and Watson (2007).

We remark important instabilities over time and between methods. Firstly, the suggested number of factors varies significantly across the criteria–in the full sample case, at the end of the expanding window, it goes from 1 to 7. Typically, the estimates of the number of dynamic factors are smaller than those of static factors. Secondly, there is lot of instability over time. For example, consider the Amengual and Watson (2007) criteria in the rolling window panel. The suggested number of factors during the 80s was stable at 3, but then rose to 4 and 5 until the 2008-09 recession. A similar behavior is observed in the expanding window exercise.

This figure presents the selection of the number of factors during 1980-2012 period. The first column presents results computed for a rolling window with 192 months in size (the initial period is 1964M01–1979M12). The second column presents results for the expanding window where the time series size grows every period. AH2013 stands for Ahn and Horenstein (2013), HL2007 for Hallin and Liska (2007), O2010 for Onatski (2010), AW2007 for Amengual and Watson (2007), BN2002 and BN2007 for Bai and Ng (2002,2007), respectively. Shaded areas represent the NBER recession periods.

3.1.1 *Interpretation of Factors*

It is well known that the factors are identified up to a rotation. The estimation of F_t by principal components of X_t specifies a particular rotation matrix such that factors are orthonormal and $\Lambda'\Lambda$ is diagonal[2]. However, after the estimation, it is common practice to verify which type of variables loads on each factor. Since we have found that the number of factors is likely to change over time, it is interesting to see if their interpretation remains stable.

2. See Bai and Ng (2013) for more details on identification issues within principal components estimation of factor models

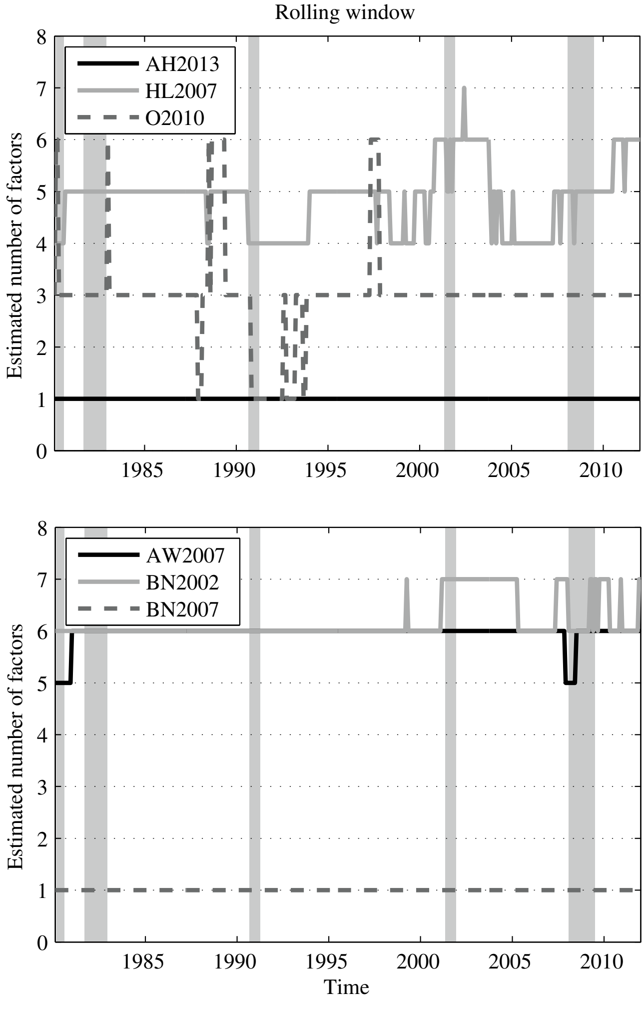
The interpretation of factors is formulated in terms of the marginal R^2 of each element in F_t for all series in X_t. To fix the ideas, we evaluate separately the part of the variance of each series explained by every factor. Then, we order the series by highest marginal R^2 s for each factor. We start with the initial period 1964M01-1979M12 and expand the panel recursively month by month. The results are presented in Figures 3 and 4. Consider, for example, the first north-west panel in Figure 3. The blue line corresponds to the highest marginal R^2 of the first factor, regardless the series. The five series in the text box are those that load the most on F_1 during that period, in descending order. This exercise reveals that the variation in the growth rate of the industrial production index of manufacturing industries (IP: mfg) is explained by more than 82 % by the first factor during 1980, but its explanatory power decreases to 77 % for the full sample. We note that the interpretation of the first factor did not change over time, it is highly related to the real sector, which includes the other series like employment and capital utilization. The explanatory power of the second factor did not change either: it represents the credit spread and the long term spread measures. Its determination coefficient goes from 65 % during the 1980 decade to 55 % at the end of the sample.

On the other hand, the interpretations of the third and the fourth factor have changed through the sample. The vertical lines correspond to periods where the ordering of five most-explained series by the factor has changed. Before 1986, F_3 was clearly related to the term structure of interest rate, but subsequently became an inflation factor. In addition, its explanatory power has risen significantly over time, especially from the year 1990. On the contrary, the fourth factor was related to prices before 1990 and then became a term structure factor from 1999 onwards. The results for factors five to eight are presented in Figure 4. The fifth factor's interpretation remained quite stable over time–it explains around 35 % of the variation in short term spreads. The sixth factor exhibited an interesting behavior. It is clearly related to the stock markets, with a respectable R^2 of 30 % for the S&P industrial returns until 1990. However, between 2001 and 2008, it explained almost 60 % of the variation in total reserves growth–clearly making it a monetary factor. Finally, the interpretations of F_7 and F_8 have evolved a lot during the 1980-2011 period. In the case of F_7, it changes from being an exchange rates, inflation, and stock market factor to a housing market factor in 2001.

Now, let us see how the estimated number of factors relates to their interpretations. Consider, for example, the Bai-Ng (2002) criteria at the south-east panel in Figure 1. The estimated number of factors is five until 1984M02. Hence, the underlying states of the economy until 1984 were: real, credit spread, term structure of interest rates, inflation, and term spread. Then, from 1984 to 1991, the estimated K grows to six, implying the following decomposition of elements of F_t : real, credit spread, inflation and term structure, term spread, and stock market factors. The ordering is important since F_t is estimated by principal components: they are ordered by explanatory power of the total variance of X_t . Between 1992 and 2001, a seventh factor is suggested by the information criteria. The interpretation from the previous period did not change except that the seventh factor is also related to

FIGURE 2

Number of Factors over Time: Financial Panel

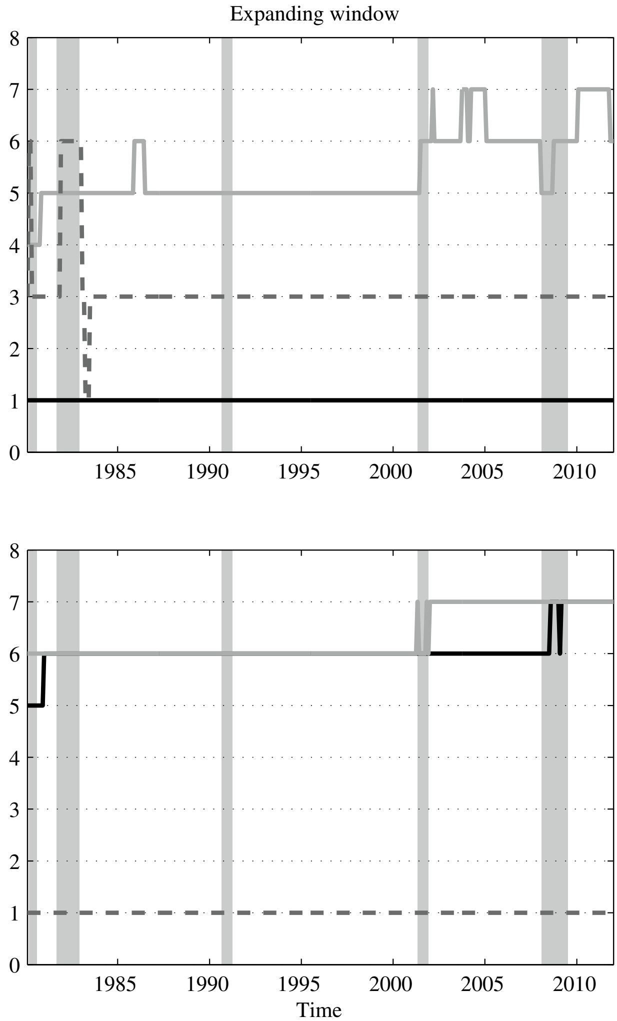


the stock market. Between 2002 and mid-2009, eight factors are needed. Now, the 6th, 7th, and 8th components correspond to monetary aggregates, housing market, and stock market, respectively. Finally, when we consider the full sample, seven factors are estimated.

3.2 *Number of Factors in a Large Panel of Financial Variables*

As noted by Onatski (2012), macroeconomic panels may suffer from a weak factor structure. In fact, macroeconomic aggregates and sectoral data are strongly

FIGURE 2 (CONTINUED)



NOTE: This figure presents the selection of the number of factors during 1980-2012 period. The first column presents results computed for a rolling window with 240 months in size (the initial period is 1960M01 - 1979M12). The second column presents results for the expanding window where the time series size grows every period. Shaded areas represent the NBER recession periods.

correlated within groups but less across them. For example, inflation series are very similar amongst each other but much less correlated to employment indicators. The presence of correlation clusters may alter the strength of the *common* factor structure and hence the estimation of pervasive factors.

The factor analysis has been applied in finance to characterize the determinants of a large set of returns. In this section, we consider a large financial data set from Jurado, Ludvigson, and Ng (2013), which is an update from Ludvigson and Ng (2007). There are 147 financial market variables observed from 1960M01 to 2011M12. The data description is available in Appendix B. Figure 2 shows the selection of the number of factors over time. There is less instability in case of the information criteria on second row panels for both rolling and expanding windows, in comparison to results from the macroeconomic panel in Figure 1. Interestingly, Hallin and Liska (2007) and Onatski (2010) suggest that number of factors varies more, especially for the rolling window. The former seems to be highly unstable since the late 90s while the latter suggests between 1 and 6 factors during 1988-1998 period.

In the Appendix, we presented more examples with other US and Canadian macroeconomic panels.. Overall, using a battery of selection methods, we find robust evidence that the number of factors is changing over time. One can offer several explanations for this finding. Possibly, if the factors represent the latent states of economy, these could be more or less pervasive over time such that their number is harder to estimate. The structural changes, such as adoption of new monetary and fiscal policies, can also affect the way observable series load on the factors. This hypothesis of structural instability can be represented by time-varying factor loadings. Hence, the number of factors is always the same but a subset of them may become more or less related to the series in X_t.

This figure presents the selection of the number of factors during 1980-2012 period. The first column presents results computed for a rolling window with 240 months in size (the initial period is 1960M01–1979M12). The second column presents results for the expanding window where the time series size grows every period. Shaded areas represent the NBER recession periods.

In the next section we will investigate if the structural instability in factor loadings can alter the selection of the number of factors. Both exact and approximate factor models will be considered.

4. Monte Carlo Simulation Exercise I: Time-varying Factor Loadings

The aim of this simulation is to assess the robustness of different tests and information criteria used when selecting the number of factors in a static factor model. Recall the model:

$$\mathbf{X_t} = \wedge_\mathbf{t} \mathbf{F_t} + \mathbf{e_t},$$

$$\wedge_t = \wedge_0 + h_{NT} \zeta_t .$$

The focus is on instabilities of the factor loadings, $\wedge_t$; hence, one needs to impose a stochastic process for them. We consider several cases that can be summarized as follows (see BPSW 2012):

Case 1: The factor loadings do not vary over time.

- $q = 2$,
- $(N,T) \in \{(50,100),(100,200)\}$,
- $\Lambda_t = \Lambda_0, \forall_t$,
- $\lambda_{i,j} \sim N(0,1)$, $F_{j,t} \sim N(0,1)$, $e_{i,t} \sim N(0,1)$, $i = 1,,N$, $j = 1,,q$ and $t = 1,,T$.

Case 2: The factor loadings are random variables.

- $q = 2$,
- $(N,T) \in \{(50,100),(100,200)\}$,
- $F_{j,t} \sim N(0,1)$, $e_{i,t} \sim N(0,1)$,
- $\Lambda_t = \zeta_t$.

For each $t \in \{1...T\}$ we draw $\lambda_{i,j}$, $1 \le j \le N$ and $1 \le i \le q$ from $N(0,1)$ distribution.

Case 3: Single large deterministic break with $h_{NT} = 10$.

- $q = 2$,
- $(N,T) \in \{(50,100),(100,200)\}$,
- $\lambda_{i,j} \sim N(0,1)$, $F_{j,t} \sim N(0,1)$, $e_{i,t} \sim N(0,1)$,
- $\Lambda_t = \begin{cases} \Lambda_0 & \textit{for } t = 1,...,T/2 \\ \Lambda_t = \Lambda_0 + 10\,\Lambda_0 & \textit{for } t > T/2 \end{cases}$.

Case 4: Single small deterministic break with $h_{NT} = 1$.

- $q = 2$,
- $(N,T) \in \{(50,100),(100,200)\}$,
- $\lambda_{i,j} \sim N(0,1)$, $F_{j,t} \sim N(0,1)$, $e_{i,t} \sim N(0,1)$,
- $\Lambda_t = \begin{cases} \Lambda_0 & \textit{for } t = 1,...,T/2 \\ \Lambda_t = \Lambda_0 + \Lambda_0 & \textit{for } t > T/2 \end{cases}$.

Case 5: Random walk

- $q = 2$,
- $(N,T) \in \{(50,100),(100,200)\}$,
- $F_{j,t} \sim N(0,1)$, $e_{i,t} \sim N(0,1)$,
- $\Lambda_t = \Lambda_{t-1} + \zeta_t$ where $\left(\zeta_{i,j}\right)_t \sim N(0,1)$. The sequence Λ_t at $t = 0$ is initialized from $\lambda_{i,j} \sim N(0,1)$.

To complete the simulation exercise, we also consider several degrees of cross-sectional and time dependence among idiosyncratic components in the observational equation, e_t. In particular, we follow Boivin and Ng (2005), Onatski (2012), and Dufour and Stevanovic (2013). Assuming that

$$e_{i,t} = \rho_N e_{i-1,t} + \zeta_{i,t}$$

and

$$\zeta_{i,t} = \rho_T \zeta_{i,t-1} + \varepsilon_{i,t}$$

$$\varepsilon_{i,t} \sim N(0,1).$$

Hence, the parameter ρ_N drives the degree of cross-sectional dependence while ρ_T is responsible for serial correlation among e_t. For each factor loadings cases above, following Dufour and Stevanovic (2013), we consider four correlation structures of e_t:

- *Exact factor structure*: $\rho_N = 0$ and $\rho_T = 0$,
- *Cross-sectional dependence:* $\rho_N = 0.5$ and $\rho_T = 0$,
- *Serial correlation*: $\rho_N = 0$ and $\rho_T = 0.9$,
- *Cross-sectional and serial dependence*: $\rho_N = 0.5$ and $\rho_T = 0.9$.

In addition, we consider two sets of panel dimensions: $N = 50, T = 100$ (small sample) and $N = 100, T = 200$ (large sample).

The Monte Carlo exercise consists of simulating 1000 times for each case, small and large samples, each correlation structure, and then applying all tests or criteria. For each selection procedure, we compute the percentage of underestimation, overestimation, and exact estimation. The mean and standard deviation of estimated number of factors are also computed.

4.1 *Results and Discussion*

The results are summarized in four tables. Table 1 shows the simulation results in the case of exact factor structure. Table 2 presents the performance of selection methods in presence of cross-sectional dependence, while Table 3 presents results where only univariate serial correlation of e_t is considered. Lastly, Table 4 shows the behavior of selection methods in the case of the weakest factor structure implied by the presence of both cross-sectional and serial dependence.

Overall, it is most problematic when the factor loadings follow a random walk (case 5). In that case, each test and information criteria fails to capture the true number of factors in both small and large samples and in all four correlation structures of the idiosyncratic component (see Tables 1–4). In particular, Ahn and Horenstein (2013) and Bai and Ng (2007) systematically underestimate the number of factors, while the others largely overestimate.

In the case of classical factor structure, $\rho_N = 0$ and $\rho_T = 0$, results summarized in Table 1 show that, having a break on loading factors (cases 3 and 4) does not prevent the identification of the good number of factors in the large sample; however, this is not always the case in the small sample. For example, when there is a high break on loading factors (case 3), Hallin and Liska (2007) underestimate the number of factors at least 16 % of the time, while Amengual and Watson (2007) information criteria overestimate approximately 15 %. Hallin and Liska (2007) have the worst record on estimating the true number of factors in a small sample and this may be due to less accuracy in estimating a spectral density with a small amount of data points.

However, as soon as we allow for time and/or cross-sectional dependence, the amplitude of the break increases the probability to undercover the true number of factors. For example, Bai and Ng (2002) information criteria is a perfect estimator when there were no dependence, but overestimates q when allowing for a break. Moreover, as the break becomes larger, the $\hat{q}$ also becomes larger: Table 2 shows that, in the large sample, when the break is 1 the estimated number of factors is three, versus eight when the amplitude of the break is 10. BPSW and Chen, Dolado, and Gonzalo (2014) find similar behavior of the Bai and Ng (2002) information criteria IC_{p2}. Another observation in Table 2 concerns Hallin and Liska (2007). In the large sample, it usually overestimates the number of factors when the magnitude of the break is 10 but performs perfectly when the break is smaller.

Strong time dependence leads many tests and information criteria to fail in identifying q even in the case of constant factor loadings. As expected, the situation is worse in small samples. However, even when the panel dimensions are larger, only Ahn and Horenstein (2013) and Bai and Ng (2007) perform well.

To summarize, the results from this extensive simulation exercise show that structural instabilities, taking several forms of time-variant factor loadings, together with cross-sectional and time dependence of the idiosyncratic component, do alter the estimation of the number of factors across many popular selection methods used in the literature.

Consequences

Here we discuss several consequences of the previous results for empirical analysis. Diffusion indices have been very popular in forecasting within the factor-augmented regressions. The typical framework consists of the forecasting equation for a series of interest y_t:

$$y_{t+h} = \alpha + \rho y_t + \beta F_t + \xi_{t+h}, \tag{9}$$

where a large number of potential predictors obey a factor model

$$X_t = \Lambda F_t + e_t.$$

Hence, the question is how the forecasting performance is affected in the presence of irregularities in the observational equation. Chen, Dolado, and Gonzalo (2014)

TABLE 1

MC SIMULATIONS: FACTOR LOADINGS INSTABILITIES WITH EXACT FACTOR STRUCTURE

1*	(N=50; T=100)					(N=100; T=200)				
Ahn and Horenstein (2013)										
Case	1	2	3	4	5	1	2	3	4	5
Under	0.5	0.6	0	0	100	0	0	0	0	100
Over	0	0	0	0	0	0	0	0	0	0
Average	1.995	1.994	2	2	1	2	2	2	2	1
Std	0.0706	0.0773	0	0	0	0	0	0	0	0
Hallin and Liska (2007)										
Case	1	2	3	4	5	1	2	3	4	5
Under	45.5	27.9	20.2	16.90	0	0	0	0	0	0
Over	0	0	0.9	0	100	0	0	0	0	100
Average	1.537	1.692	1.605	1.68	7.96	2	2	2	2	7.998
Std	0.5147	0.521	1.81	0.71	0.2011	0	0	0	0	0.447
Onatski (2010)										
Case	1	2	3	4	5	1	2	3	4	5
Under	0	0	0	0	0	0	0	0	0	0
Over	1	0.8	1.10	0.8	100	0.6	0.9	1.3	0.6	100
Average	2.018	2.013	2.014	2.009	7.997	2.006	2.016	2.013	2.011	7.998
Std	0.24	0.2071	0.1607	0.1045	0.0547	0.0773	0.2276	0.1133	0.2023	0.447
Bai and Ng (2007)										
Case	1	2	3	4	5	1	2	3	4	5
Under	0.3	0.3	1.3	2	100	0	0	0	0	100
Over	16.5	13.1	0	0.1	0	0	0.1	0	0	0
Average	2.166	2.132	1.987	1.98	1	2	2.001	2	2	1
Std	0.3906	0.3588	0.11	0.14	0	0	0.0316	0	0	0
Bai and Ng (2002)										
Case	1	2	3	4	5	1	2	3	4	5
Under	0	0	0	0	0	0	0	0	0	0
Over	0	0	0	0	100	0	0	0	0	100
Average	2	2	2	2	8	2	2	2	2	8
Std	0	0	0	0	0	0	0	0	0	0

TABLE 1 (CONTINUED)

1*	(N=50; T=100)					(N=100; T=200)				
Amengual and Watson (2007)										
Case	1	2	3	4	5	1	2	3	4	5
Under	1.4	2	0	0	0	0	0	0	0	0
Over	0	0	15.8	0	100	0	0	0	0	100
Average	1.986	1.979	2.213	2	6	2	2	2	2	6
Std	0.1175	0.1503	0.5639	0	0	0	0	0	0	0

NOTE: This table presents the selection of the number of factors with factor loadings instabilities without any dependencies within idiosyncratic components. Case 1: constant factor loadings. Case 2: factor loadings are random variables. Case 3: single large deterministic break on loadings. Case 4: single small deterministic break on loadings. Case 5: each factor loading follows a random walk.

TABLE 2

MC SIMULATIONS: FACTOR LOADINGS INSTABILITIES WITH CROSS-SECTIONAL DEPENDANCE

1*	(N=50; T=100)					(N=100; T=200)				
Ahn and Horenstein (2013)										
Case	1	2	3	4	5	1	2	3	4	5
Under	0	0	0	0	100	0	0	0	0	100
Over	0	0	0	0	0	0	0	0	0	0
Average	2	2	2	2	1	2	2	2	2	1
Std	0	0	0	0	0	0	0	0	0	0
Hallin and Liska (2007)										
Case	1	2	3	4	5	1	2	3	4	5
Under	0	0	0	0	0	0	0	0	0	0
Over	0	0	40.2	0.1	100	0	0	57.4	0	100
Average	2	2	2.685	2.001	7.999	2	2	2.768	2	8
Std	0	0	1.06	0.0316	0.0316	0	0	0.8372	0	0
Onatski (2010)										
Case	1	2	3	4	5	1	2	3	4	5
Under	0	0	0	0	0	0	0	0	0	0
Over	2.5	3.2	6.8	3.7	100	1.1	1.9	0.7	2	100
Average	2.045	2.043	2.277	2.231	8	2.011	2.028	2.009	2.029	8
Std	0.3565	0.2593	1.0571	0.7671	0	0.1044	0.2435	0.1222	0.2573	0

TABLE 2 (CONTINUED)

1*	(N=50; T=100)					(N=100; T=200)				
Bai and Ng (2007)										
Case	1	2	3	4	5	1	2	3	4	5
Under	2.8	5.6	1	1.6	100	0	0	0	0	100
Over	0	0	0	0	0	0	0	0	0	0
Average	1.972	1.944	1.99	1.984	1	2	2	2	2	1
Std	0.1651	0.23	0.0995	0.1255	0	0	0	0	0	0
Bai and Ng (2002)										
Case	1	2	3	4	5	1	2	3	4	5
Under	0	0	0	0	0	0	0	0	0	0
Over	100	100	100	100	100	0	0	100	100	100
Average	3	3	8	8	8	2	2	8	3	8
Std	0	0	0	0	0	0	0	0	0	0
Amengual and Watson (2007)										
Case	1	2	3	4	5	1	2	3	4	5
Under	0	0	0	0	0	0	0	0	0	0
Over	91.40	93.8	100	100	100	33.8	64.9	100	96.1	100
Average	4.603	4.939	6	5.959	6	2.42	3.099	6	4.578	6
Std	1.3263	1.2739	0	0.2267	0	0.6709	1.0697	0	1.1914	0

NOTE: This table presents the selection of the number of factors with factor loadings instabilities with cross-sectional dependencies within idiosyncratic components. Case 1: constant factor loadings. Case 2: factor loadings are random variables. Case 3: single large deterministic break on loadings. Case 4: single small deterministic break on loadings. Case 5: each factor loading follows a random walk.

TABLE 3

MC SIMULATIONS: FACTOR LOADINGS INSTABILITIES WITH TIME DEPENDENCE

1*	(N=50; T=100)					(N=100; T=200)				
Ahn and Horenstein (2013)										
Case	1	2	3	4	5	1	2	3	4	5
Under	98.8	85.6	0	0.1	100	1.7	1.8	0	0	100
Over	1.2	13.9	100	99.80	0	98.3	98.2	100	100	0
Average	1.024	1.283	3	2.999	1	2.966	2.964	3	3	1
Std	0.2179	0.6938	0	0.0837	0	0.2587	0.266	0	0	0

TABLE 3 (CONTINUED)

1*	(N=50; T=100)					(N=100; T=200)				
Hallin and Liska (2007)										
Case	1	2	3	4	5	1	2	3	4	5
Under	0	0	0.6	0	0	0	0	0	0	0
Over	100	100	99.4	100	100	100	100	100	100	100
Average	3.066	3.051	3.436	3.172	7.789	3.051	3.062	3.639	3.157	7.912
Std	0.3432	0.3106	0.9863	0.541	0.4782	0.3413	0.4151	1.0667	0.6187	0.3104
Onatski (2010)										
Case	1	2	3	4	5	1	2	3	4	5
Under	6.6	5.3	0	0	0	0	0	0	0	0
Over	90.8	92.8	100	100	100	100	100	100	100	100
Average	3.262	3.259	3.472	3.373	7.906	3.0360	3.032	3.043	3.033	7.862
Std	1.2661	1.1838	1.0965	1.0232	0.3021	0.3699	0.2917	0.3182	0.3162	0.348
Bai and Ng (2007)										
Case	1	2	3	4	5	1	2	3	4	5
Under	15.9	12.3	5.2	15.5	100	0	0	0	0	100
Over	0	0	0	0	0	0	0	0	0	0
Average	1.841	1.877	1.948	1.845	1	2	2	2	2	1
Std	0.3659	0.3286	0.2221	0.3621	0	0	0	0	0	0
Bai and Ng (2002)										
Case	1	2	3	4	5	1	2	3	4	5
Under	0	0	0	0	0	0	0	0	0	0
Over	100	100	100	100	100	100	100	100	100	100
Average	8	8	8	8	8	8	8	8	8	8
Std	0	0	0	0	0	0	0	0	0	0
Amengual and Watson (2007)										
Case	1	2	3	4	5	1	2	3	4	5
Under	0	0	0	0	0	0	0	0	0	0
Over	100	100	100	100	100	100	100	100	100	100
Average	5.139	5.246	5.939	5.504	6	4.9710	6	6	5.383	6
Std	1.0596	1.0414	0.3121	0.8711	0	1.1625	0	0	0.966	0

NOTE: This table presents the selection of the number of factors with factor loadings instabilities with serial dependencies within idiosyncratic components. Case 1: constant factor loadings. Case 2: factor loadings are random variables. Case 3: single large deterministic break on loadings. Case 4: single small deterministic break on loadings. Case 5: each factor loading follows a random walk.

TABLE 4

MC SIMULATIONS: FACTOR LOADINGS INSTABILITIES WITH TIME AND CROSS-SECTIONAL DEPENDENCES

1*	(N=50; T=100)					(N=100; T=200)				
Ahn and Horenstein (2013)										
Case	1	2	3	4	5	1	2	3	4	5
Under	100	100	0	99.7	100	100	100	0	27.6	100
Over	0	0	100	0.3	0	0	0	100	72.4	0
Average	1	1	3	1.006	1	1	1	3	2.448	1
Std	0	0	0	0.1094	0	0	0	0	0.8945	0
Hallin and Liska (2007)										
Case	1	2	3	4	5	1	2	3	4	5
Under	0.5	0.1	8.3	1.6	0	0	0	0	0	0
Over	98.4	99.8	91.7	98.30	100	100	100	100	100	100
Average	3.009	3.042	3.268	3.104	7.782	3.063	3.052	3.838	3.163	7.933
Std	0.2811	0.323	1.3386	0.5844	0.4676	0.3862	0.3456	1.1752	0.5889	0.2731
Onatski (2010)										
Case	1	2	3	4	5	1	2	3	4	5
Under	24.8	15.5	0	0.2	0	0	0.2	0	0	0
Over	67.5	78.9	100	99.6	100	100	99.8	100	100	100
Average	2.934	3.206	3.702	3.543	7.877	3.052	3.07	3.124	3.075	7.937
Std	1.5748	1.5277	1.2177	1.124	0.3376	0.3089	0.4138	0.5088	0.3867	0.2472
Bai and Ng (2007)										
Case	1	2	3	4	5	1	2	3	4	5
Under	19.3	16.4	10.7	28.7	100	0	0	0	0	100
Over	0	0	0	0	0	0	0	0	0	0
Average	1.807	1.836	1.893	1.713	1	2	2	2	2	1
Std	0.3949	0.3705	0.3093	0.4526	0	0	0	0	0	0
Bai and Ng (2002)										
Case	1	2	3	4	5	1	2	3	4	5
Under	0	0	0	0	0	0	0	0	0	0
Over	100	100	100	100	100	100	100	100	100	100
Average	8	8	8	8	8	8	8	8	8	8
Std	0	0	0	0	0	0	0	0	0	0

TABLE 4 (CONTINUED)

1*	(N=50; T=100)					(N=100; T=200)				
Amengual and Watson (2007)										
Case	1	2	3	4	5	1	2	3	4	5
Under	0	0	0	0	0	0	0	0	0	0
Over	100	100	100	100	100	100	100	100	100	100
Average	5.777	5.869	5.999	5.914	6	5.405	5.983	6	5.698	6
Std	0.5635	0.4336	0.0316	0.3616	0	0.9273	0.1369	0	0.6937	0

NOTE: This table presents the selection of the number of factors with factor loadings instabilities with both cross-sectional and serial dependencies within idiosyncratic components. Case 1: constant factor loadings. Case 2: factor loadings are random variables. Case 3: single large deterministic break on loadings. Case 4: single small deterministic break on loadings. Case 5: each factor loading follows a random walk.

show, using simulations, that imposing a priori number of factors that ignores the existence of a large break on Λ can worsen the forecasting power of the factor-augmented regressions. Overestimating the number of factors can help, but this entails more estimation uncertainty that ultimately increases the mean squared predicted errors. Barhoumi, Darné, and Ferrara (2013) compare several selection methods in the pseudo out-of-sample forecasting exercise and find that setting the number of factors with the Alessi, Barigozzi and Capasso (2010) information criterion (a modification of Bai and Ng, 2002) produces significantly lower squared prediction errors.

Our results contribute to these findings by showing that many selection methods typically overestimate the number of factors. Hence, if they are used to assess the dimension of F_t to include in (9), a similar forecasting behavior is expected to occur. More importantly, we showed that there are cases where Ahn-Horenstein (2013) and Bai and Ng (2007) tests underestimate the true number of latent common components. Obviously, this will misspecify the forecasting equation (9) as some important predictors would be omitted.

Another area of interest for factor models is the structural analysis. Since Bernanke, Boivin, and Eliasz (2005), the factor-augmented VAR (FAVAR) approach has been heavily used to identify and estimate the effects of structural shocks (monetary, news, productivity, credit, etc.) on real economy of many countries. The FAVAR model consists of the state-space representation

$$X_t = \Lambda F_t + e_t, \tag{10}$$

$$F_t = \Phi(L) F_{t-i} + u_t. \tag{11}$$

where u_t are the reduced-form disturbances related to the structural shocks via $u_t = H\varepsilon_t$. The objects of interest are impulse responses of X_t to the structural shocks ε_t

$$X_t = [I - \Phi(L)L]^{-1} H\varepsilon_t. \tag{12}$$

Clearly, the misspecification, and particularly the underestimation, of the number of elements in F_t will alter both the identification of structural shocks and the estimation of the impulse responses. An extensive study on the consequences on forecasting and structural analysis goes beyond the scope of this paper, but is a part of our research agenda.

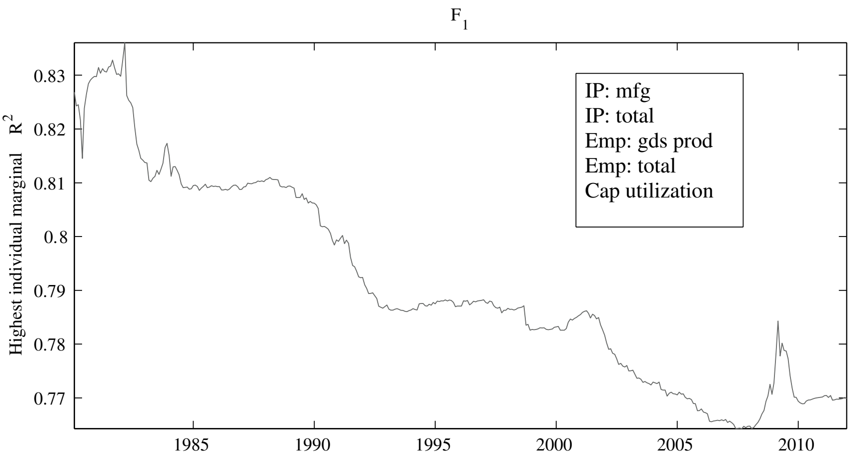
Conclusion

The objective of this paper is to verify the robustness of most important selection methods to identify the number of factors in large data sets. Empirically, we show that, in both large macroeconomic and financial panels, the estimated number of factors varies significantly across time procedures. To provide an explanation of these findings we conduct an extensive Monte Carlo simulation exercise with several time-varying processes for factor loadings in both exact and approximate factor model structures.

The simulation results show that structural instabilities do alter the estimation of the number of factors across all six most popular selection methods used in the literature. Their performance is particularly affected when factor loadings behave as random walks and in the presence of cross-sectional and time dependencies across idiosyncratic components. More research is needed to explore the exact theoretical reasons for the systematic failure of these procedures. In addition, we hope this work will provide a basis for pursuing research on developing new estimators of the factor space rank in the presence of time instabilities.

FIGURE 3

Marginal R^2 of Factors over Time



Note: This figure presents the marginal R^2 of first four factors from 1980 to 2012. For example, the first north-west panel shows the highest individual marginal R^2 of the first factor, F_1. Through all sample, the first factor has the highest marginal contribution to variations in the growth rate of industrial production in manufacturing sector (IP: mfg). The box contains several most explained series by the first factor, in terms of the marginal R^2.

FIGURE 3 (CONTINUED)

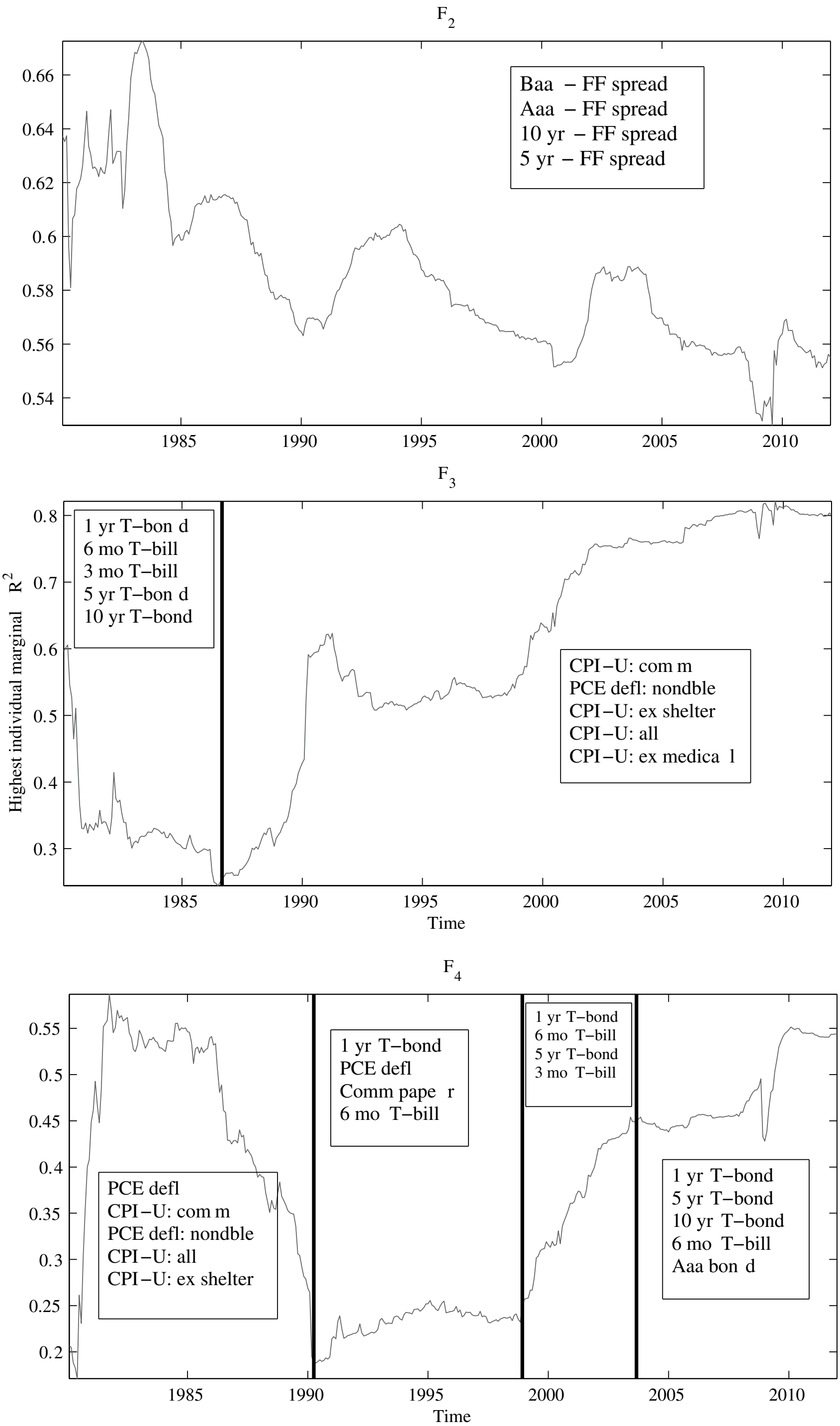


FIGURE 4

MARGINAL R^2 OF FACTORS OVER TIME

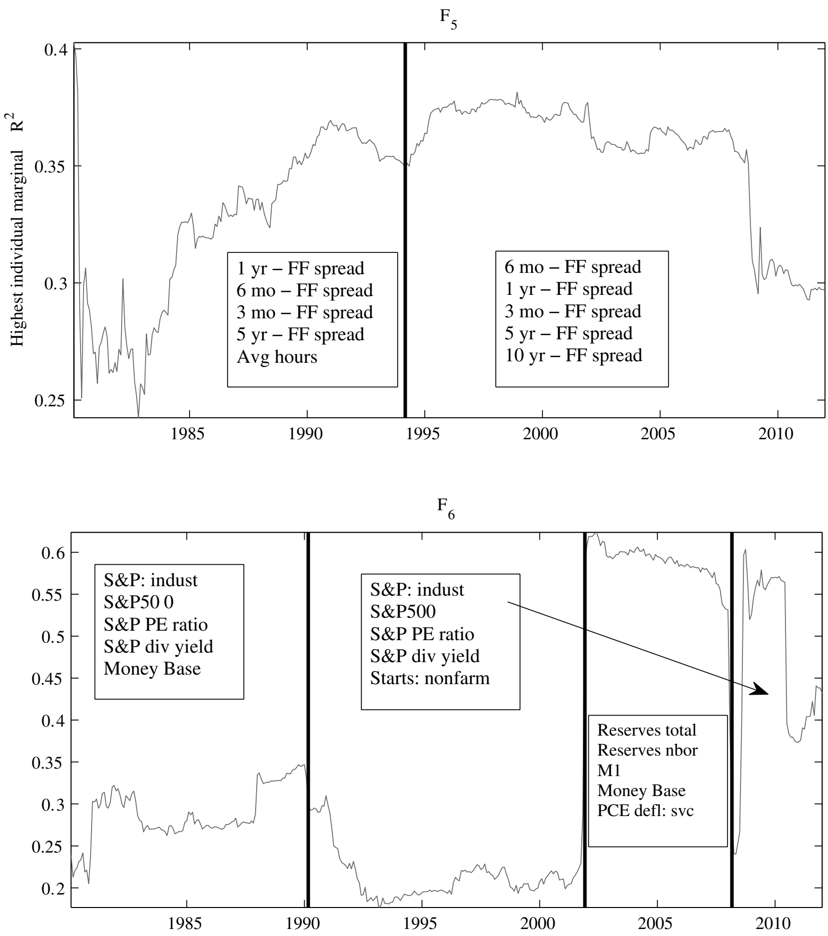
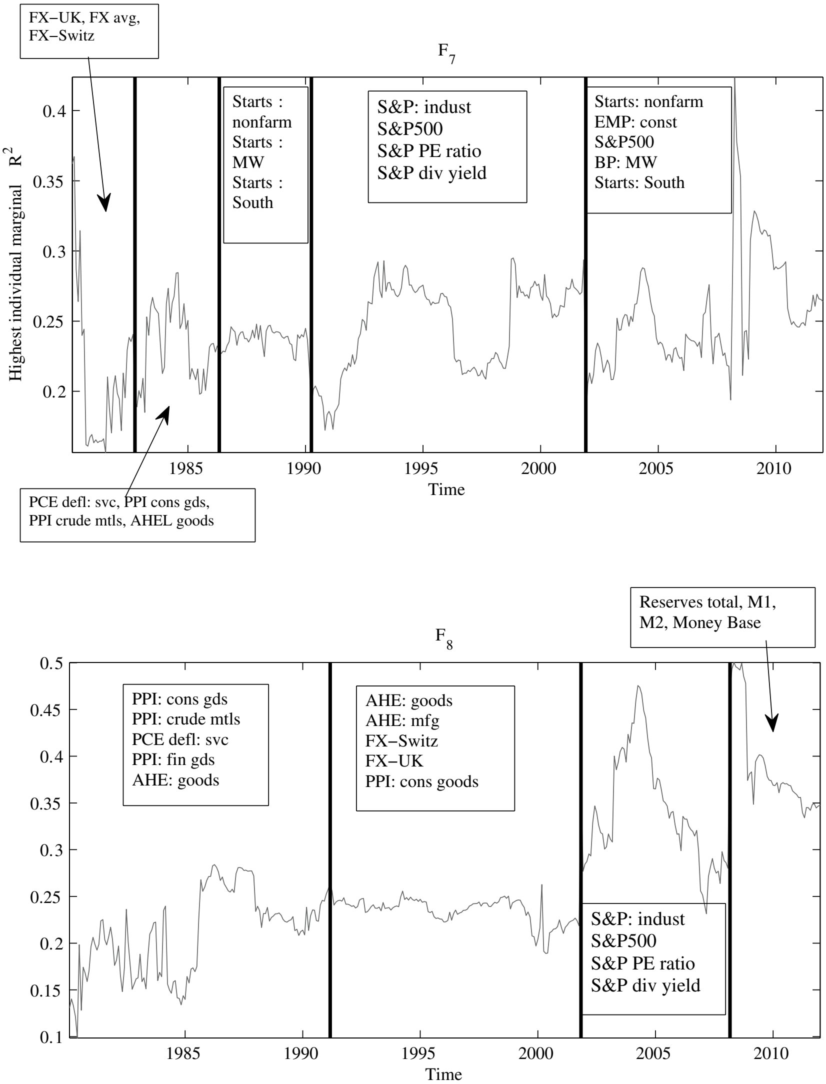


FIGURE 4 (CONTINUED)



NOTE: This figure is the continuation of Figure 3 and presents the marginal R^2 of factors five to eight from 1980 to 2012. For example, the first north-west panel shows the highest individual marginal R^2 of the first factor, F_1.

APPENDIX

Appendix A: Additional Empirical Results

Figure 5 shows the estimated number of factors over time for the macroeconomic panel used in Boivin, Giannoni and Stevanovic (2013) which is an update of the data set in Bernanke, Boivin and Eliasz (2005). There are 124 variables observed from 1959M01 to 2009M06. This panel is very similar to the one used by Jurado, Ludvigson and Ng (2013) except for the stationarity assumptions on a subset of series. In this data set, interest, unemployment and inflation rates are supposed stationary, therefore they enter X_t in levels, contrary to Jurado, Ludvigson and Ng (2013) where the same series are in first difference of logs. Compared to Figure 1, these stationarity assumptions imply more factors on average over time.

FIGURE 5

Number of Factors over Time: Macroeconomic Panel from Boivin *et al.* (2013)

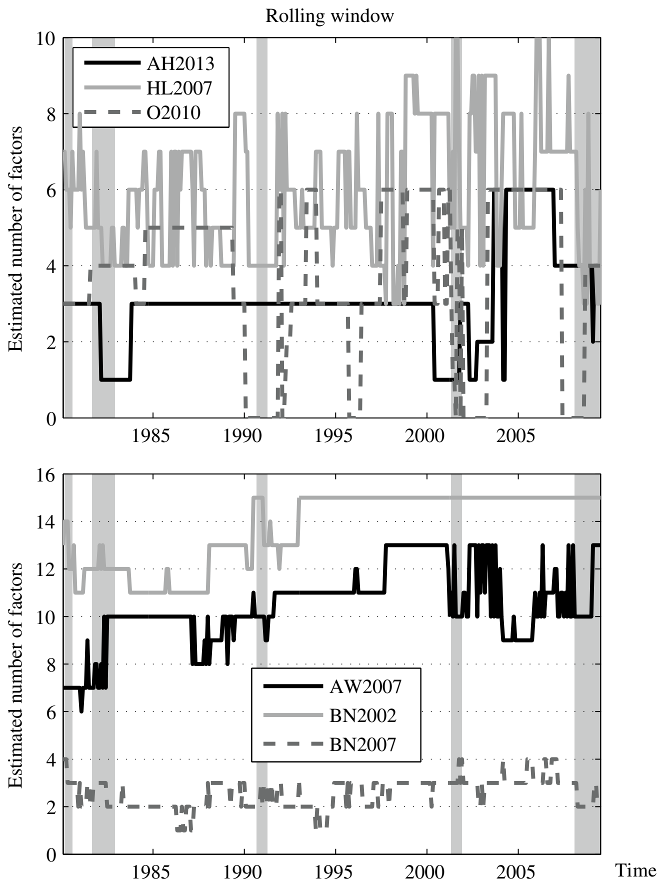
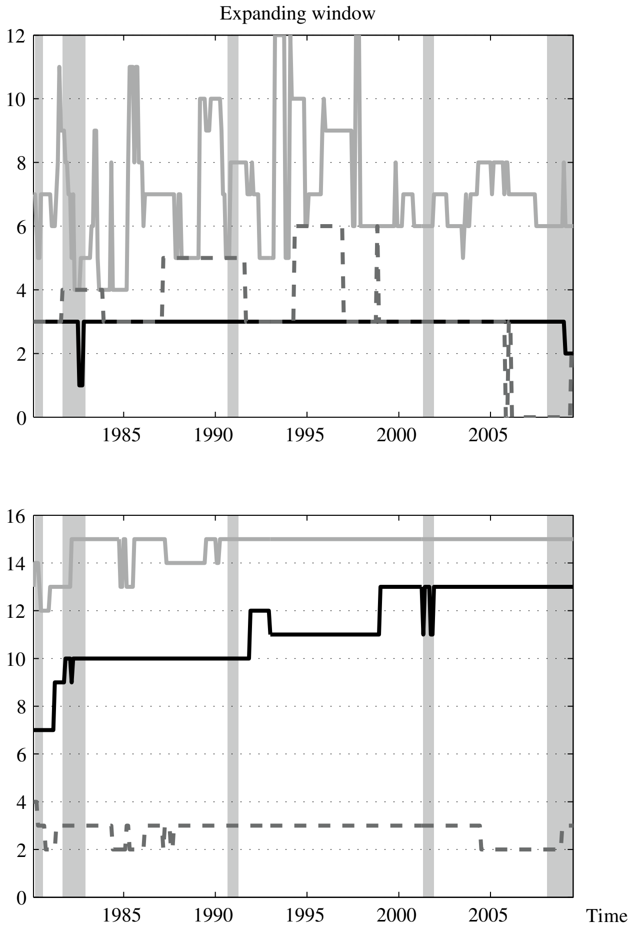


FIGURE 5 (CONTINUED)



NOTE: This figure presents the selection of the number of factors during 1980-2009 period. The first column results are computed for rolling window of size 251 months (the initial period is 1959M02–1979M12). The second column is for the expanding window where the time series size grows every period. Shaded areas represent the NBER recession periods.

FIGURE 6

NUMBER OF FACTORS OVER TIME: CANADIAN MACROECONOMIC PANEL

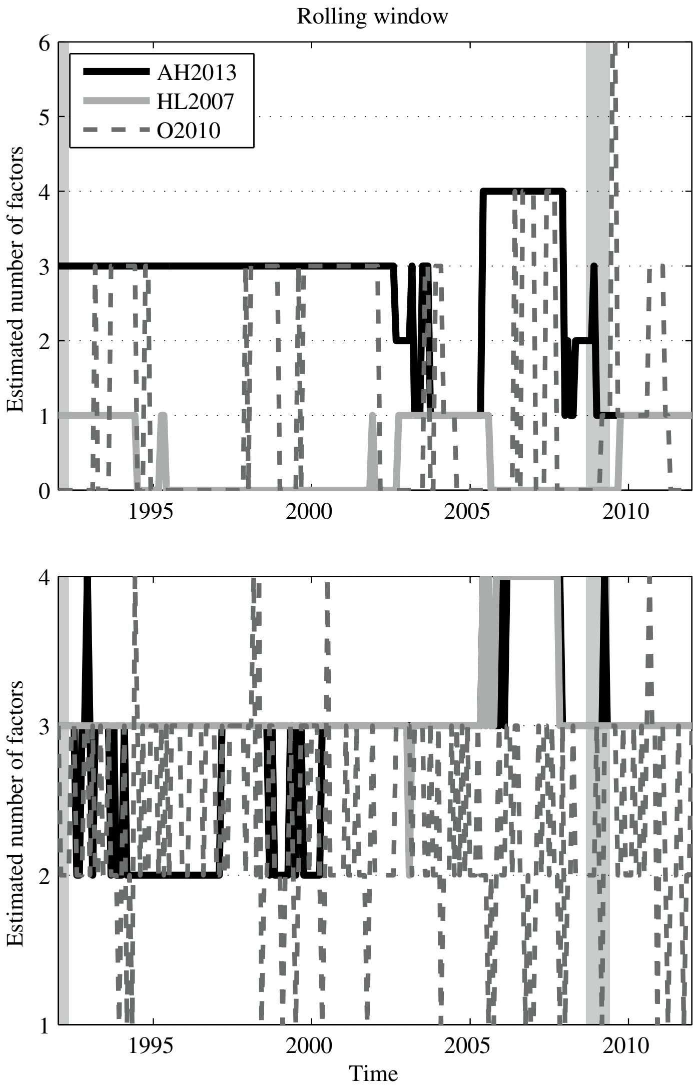
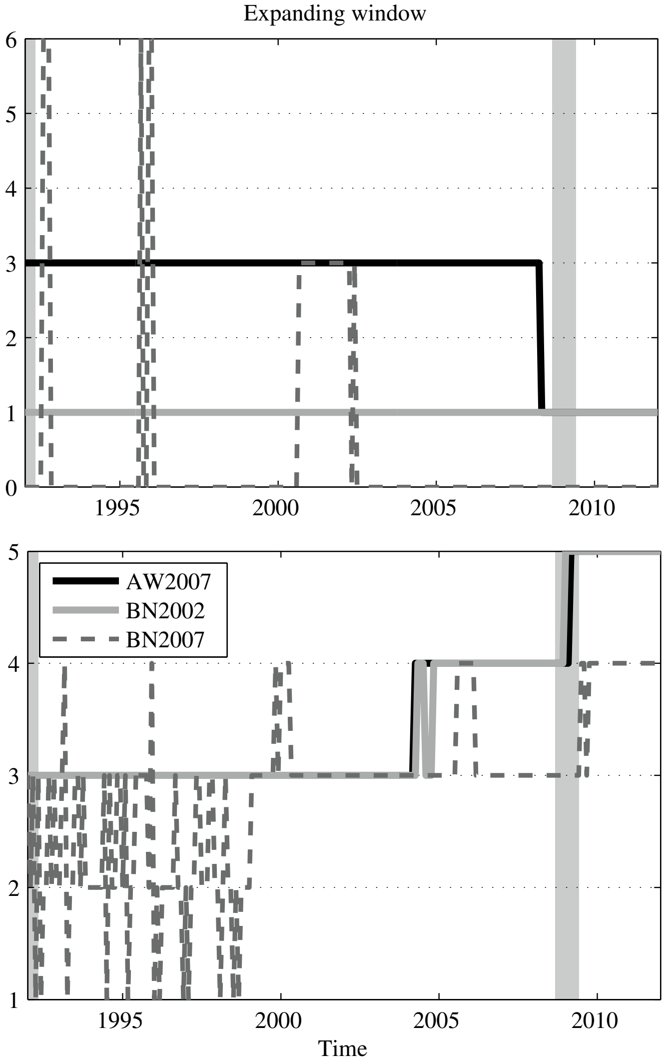


Figure 6 shows the estimated number of factors over time for a macroeconomic panel of Canadian series. The composition of the panel is very similar to the US data set used in Jurado, Ludvigson and Ng (2013). In addition, the same stationarity assumptions are imposed. There are 124 variables observed from 1981M01 to 2011M12. The Canadian macroeconomic data are typically less available and since the recent reform at Statistics Canada many series are constructed from 1981 only.

FIGURE 6 (CONTINUED)



NOTE: This figure presents the selection of the number of factors during 1992-2012 period. The first column results are computed for rolling window of size 131 months (the initial period is 1981M01–1991M12). The second column is for the expanding window where the time series size grows every period. Shaded areas represent the Canadian recession periods.

Finally, we combine the previous Canadian data set with the Jurado, Ludvigson and Ng (2013) US panel to construct a very large US-CAN macroeconomic panel containing 246 series for 1981 — 2012 period. The results are presented in Figure 7. Overall, the number of factors seems to grow over time.

FIGURE 7

Number of Factors over Time: US and Canadian Macroeconomic Panel

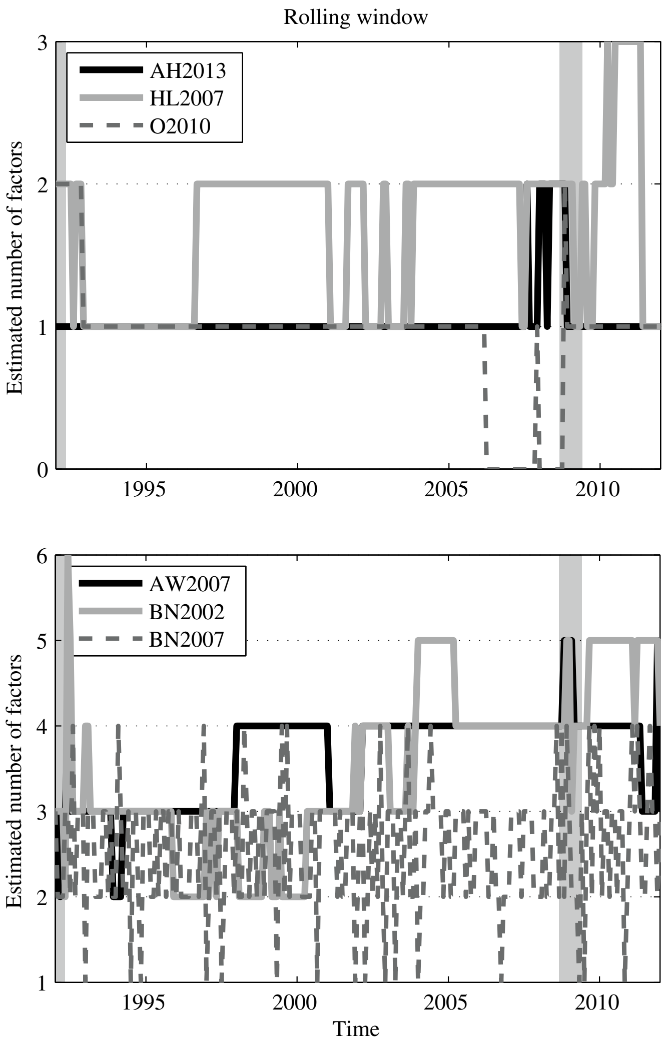
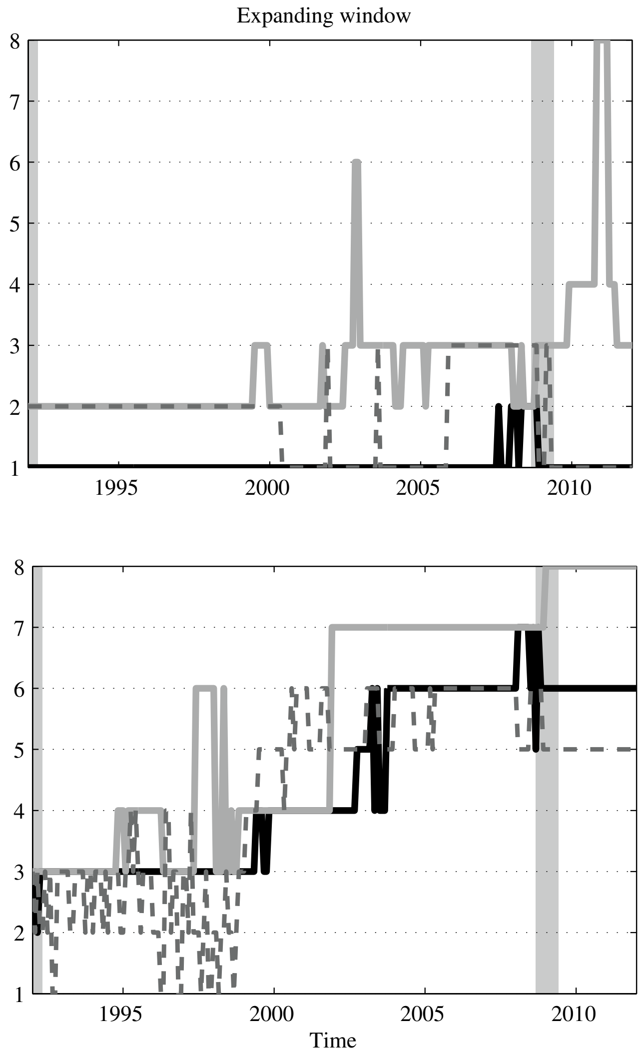


FIGURE 7 (CONTINUED)



NOTE: This figure presents the selection of the number of factors during 1992-2012 period. The first column results are computed for rolling window of size 131 months (the initial period is 1981M01–1991M12). The second column is for the expanding window where the time series size grows every period. Shaded areas represent the Canadian recession periods.

Appendix B: Data Sets

The transformation codes are: 1 — no transformation; 2 — first difference; 4 — logarithm; 5 — first difference of logarithm; 6 — second difference of logarithm. All US macro series are available from Datastream. The sources of financial data are CRSP and Kenneth French and Monika Piazzesi websites. Canadian macroeconomic series are available at StatCan.

Macroeconomic Panel from Jurado, Ludvigson and Ng (2013)

N°	Short Name	T-Code	Series description
1	PI	5	Personal income
2	PI less transfers	5	Personal income less transfers
3	Consumption	5	Real personal consumption expenditures (AC)
4	M&T sales	5	Manufacturing and trade sales
5	Retail sales	5	Sales of retail stores
6	IP: total	5	Industrial production index–total index
7	IP: products	5	Industrial production index–products, Total
8	IP: Final prod	5	Industrial production index–Final Products
9	IP: cons gds	5	Industrial Production Index–consumer goods
10	IP: cons dble	5	Industrial production index–durable consumer goods
11	IP: cons nondble	5	Industrial production index – non-durable consumer goods
12	IP: bus eqpt	5	Industrial production index–business equipment
13	IP: matls	5	Industrial production index–materials
14	IP: dble matls	5	Industrial production index–durable goods materials
15	IP: nondble matls	5	Industrial production index – non-durable goods materials
16	IP: mfg	5	Industrial production index–manufacturing
17	IP: res util	5	Industrial production index–residential utilities
18	IP: fuels	5	Industrial production index–fuels
19	NAPM prodn	1	Napm production index
20	Cap util	2	Capacity utilization
21	Help wanted indx	2	Index of help-wanted advertising (B)
22	Help wanted/unemp	2	Ratio of help-wanted ads/No. unemployed (AC)
23	Emp CPS total	5	Civilian labor force: employed, total
24	Emp CPS nonag	5	Civilian labor force: employed, non-agricultural industries

MACROECONOMIC PANEL FROM JURADO, LUDVIGSON AND NG (2013) (CONTINUED)

N°	Short Name	T-Code	Series description
25	U: mean duration	5	Unemployment by duration: average duration In weeks
26	U: mean duration	2	Unemployment by duration: average duration in weeks
27	U < 5 wks	5	Unemployment by duration: persons unemployed less than 5 weeks
28	U 5-14 wks	5	Unemployment by duration: persons unemployed 5 to 14 weeks
29	U 15+ wks	5	Unemployment by duration: persons unemployed 15 weeks +
30	U 15-26 wks	5	Unemployment by duration: persons unemployed 15 to 26 weeks
31	U 27+ wks	5	Unemployment by duration: persons unemployed 27 weeks +
32	UI claims	5	Initial claims for unemployment insurance
33	Emp: total	5	Employees on nonfarm payrolls: total private
34	Emp: gds prod	5	Employees on nonfarm payrolls–goods-producing
35	Emp: mining	5	Employees on nonfarm payrolls–mining
36	Emp: const	5	Employees on nonfarm payrolls–Construction
37	Emp: mfg	5	Employees on nonfarm payrolls–Manufacturing
38	Emp: dble gds	5	Employees on nonfarm payrolls–Durable Goods
39	Emp: nondbles	5	Employees on nonfarm payrolls–Nondurable goods
40	Emp: services	5	Employees on nonfarm payrolls–Service-Providing
41	Emp: TTU	5	Employees on nonfarm payrolls–Trade, transportation and utilities
42	Emp: wholesale	5	Employees on nonfarm payrolls–Wholesale trade.
43	Emp: retail	5	Employees on nonfarm payrolls–Retail trade
44	Emp: FIRE	5	Employees on nonfarm payrolls–Financial activities
45	Emp: Govt	5	Employees on nonfarm payrolls–Government
46	Agg wkly hours	2	Index of aggregate weekly hours (BLS)
47	Avg hrs	2	Average weekly hours of production or nonsupervisory workers private non-farm–Goods-producing

MACROECONOMIC PANEL FROM JURADO, LUDVIGSON AND NG (2013)
(CONTINUED)

N°	Short Name	T-Code	Series description
48	Overtime: mfg	2	Average weekly hours of production or nonsupervisory workers private nonfarm– Manufacturing overtime
49	Avg hrs: mfg	2	Average weekly hours, manfacturing.
50	NAPM empl	1	NAPM employment index
51	Starts: nonfarm	5	Housing starts: nonfarm(1947-58); Total farm & nonfarm(1959-)
52	Starts: NE	5	Housing starts: northeast
53	Starts: MW	5	Housing starts: midwest
54	Starts: South	5	Housing starts: South
55	Starts: West	5	Housing starts: west
56	BP: total	5	Housing authorized: total new private housing units
57	BP: NE	5	Houses authorized by build. Permits: northeast
58	BP: MW	5	Houses authorized by build. Permits: midwest
59	BP: South	5	Houses authorized by build. Permits: south
60	BP: West	5	Houses authorized by build. Permits: west
61	PMI	1	Purchasing managers index
62	NAPM new ordrs	1	Napm vendor deliveries index
63	NAPM vendor del	1	Napm vendor deliveries index
64	NAPM Invent	1	Napm inventories index
65	Orders: cons gds	5	Manufacturers new orders, Consumer goods and materials
66	Orders: dble gds	5	Manufacturers new orders, Durable goods industries
67	Orders: cap gds	5	Manufacturers new orders, Nondefense capital goods
68	Unf orders: dble	5	Manufacturers ulled orders, Durable goods industries.
69	M&T invent	5	Manufacturing and trade inventories
70	M&T invent/sales	2	Ratio, manufacturing. and trade Inventories to sales
71	M1	6	Money stock: M1
72	M2	6	Money stock: M2
73	Currency	6	Money stock: currency held by the public

Macroeconomic Panel from Jurado, Ludvigson and Ng (2013) (CONTINUED)

N°	Short Name	T-Code	Series description
74	M2 (real)	5	Money supply: real M2 (AC)
75	MB	6	Monetary base, adj. for reserve requirement changes
76	Reserves tot	6	Depository inst. reserves: total, adj. for reserve req. chgs.
77	Reserves nonbor	6	Depository inst. reserves: non-borrowed, adj. res. req. chgs.
78	C&I loans	6	Commercial and industrial loans at all commercial banks (FRED)
79	C&I loans	1	Change in commercial and industrial loans at all commercial banks (FRED)
80	Cons credit	6	Consumer credit outstanding – non-revolving
81	Inst cred/PI	2	Ratio, consumer installment credit to personal income
82	S&P 500	5	S&P?s common stock price index: composite
83	S&P: indust	5	S&P?s common stock price index: & industrials
84	S&P div yield	2	S&P?s composite common Stock: dividend yield real (S)
85	S&P PE ratio	5	S&P?s composite common stock: price-earnings ratio real (S)
86	Fed Funds	2	Interest rate: federal funds
87	Comm paper	2	3-Month AA financial commercial paper rate (FRED)
88	3 mo T-bill	2	Interest rate: U.S. Treasury bills, secondary market, 3-months.
89	6 mo T-bill	2	Interest Rate: U.S. Treasury bills, secondary market, 6-months.
90	1 yr T-bond	2	Interest rate: U.S.Treasury const. maturities,1-year.
91	5 yr T-bond	2	Interest rate: U.S.Treasury const. maturities,5-years.
92	10 yr T-bond	2	Interest rate: U.S.Treasury const. maturities,10-years.
93	Aaa bond	2	Bond yield: Moody?s Aaa corporate
94	Baa bond	1	Bond yield: Moody?s Baa corporate
95	CP-FF spread	1	CP-FF spread (AC)

MACROECONOMIC PANEL FROM JURADO, LUDVIGSON AND NG (2013)
(CONTINUED)

N°	Short Name	T-Code	Series description
96	3 mo-FF spread	1	6 mo-FF spread (AC)
97	6 mo-FF spread	1	6 mo-FF spread (AC)
98	1 yr-FF spread	1	1 yr-FF spread (AC)
99	5 yr-FF spread	1	5 yr-FF spread (AC)
100	10 yr-FF spread	1	10 yr-FF spread (AC)
101	Aaa-FF spread	1	Aaa-FF spread (AC)
102	Baa-FF spread	1	Baa-FF spread (AC)
103	Ex rate: avg	5	Nominal effective exchange rate, unit labor costs (IMF)
104	Ex rate: Switz	5	Foreign exchange rate: Switzerland–Swiss franc per U.S. $
105	Ex rate: Japan	5	Foreign exchange rate: Japan–yen per U.S. $
106	Ex rate: UK	5	Foreign exchange rate: United Kingdom–cents per pound
107	EX rate: Canada	5	Foreign exchange rate: Canada–Canadian $ per U.S. $
108	PPI: ?n gds	6	Producer price index: finished goods
109	PPI: cons gds	6	Producer price index: finished consumer goods
110	PPI: int materials	6	Producer price index: intermed mat. supplies & components
111	PPI: crudematerials	6	Producer price index: crude materials
112	Spot market price	6	Spot market price index: bls & crb: all commodities
113	PPI: nonfermaterials	6	Producer price index: nonferrous materials
114	NAPM com price	1	Napm commodity prices index
115	CPI-U: all	6	Cpi-U: all items
116	CPI-U: apparel	6	Cpi-U: apparel & upkeep
117	CPI-U: transp	6	Cpi-U: transportation
118	CPI-U: medical	6	Cpi-U: medical care
119	CPI-U: comm.	6	Cpi-U: commodities
120	CPI-U: dbles	6	Cpi-U: durables
121	CPI-U: services	6	Cpi-U: services
122	CPI-U: ex food	6	Cpi-U: all items less food

MACROECONOMIC PANEL FROM JURADO, LUDVIGSON AND NG (2013) (CONTINUED)

N°	Short Name	T-Code	Series description
123	CPI-U: ex shelter	6	Cpi-U: All items less shelter
124	CPI-U: ex med	6	Cpi-U: all items less medical care
125	PCE def	6	Pce, implicit price deflator: pce (BEA)
126	PCE defl: dlbes	6	Pce, implicit price deflator: pce durables (BEA)
127	PCE defl: nondble	6	Pce, Implicit price deflator: pce non-durables (BEA)
128	PCE defl: service	6	Pce, Implicit price deflator: pce services (BEA)
129	AHE: goods	6	Average hourly earnings of prod or non sup. workers private non-farm–Goods-Producing
130	AHE: const	6	Average hourly earnings of prod or non-sup. workers private non-farm–Construction
131	AHE: mfg	6	Average hourly earnings of prod. or non-sup. workers private non-farm–Manufacturing
132	Cons exp	2	U. Of Mich. Index of consumer expectations (UM)

FINANCIAL SERIES FROM JURADO, LUDVIGSON AND NG (2013)

N°	Short name	Name source	T-code	Series description
1	D—log(DIV)	CRSP	5	dlogD see additional details below
2	D—log(P)	CRSP	5	dlogP see additional details below
3	D—DIVreinvest	CRSP	5	dlogDre, see additional details below
4	D—Preinvest	CRSP	5	dlogPre, see additional details below
5	d-p	CRSP	4	dlog(D)-logP see additional details below
6	R15-R11	Kenneth French	1	(Small, & Hig) minus (Small, & Low) sorted on (size, & book-to-market)
7	CP	Monika Piazzesi	1	Cochrane-Piazzesi factor (Cochrane and Piazzesi, 2005)
8	Mkt-RF	Kenneth French	1	Market excess return
9	SMB	Kenneth French	1	Small Minus Big, sorted on size
10	HML	Kenneth French	1	High Minus Low, sorted on book-to-market

FINANCIAL SERIES FROM JURADO, LUDVIGSON AND NG (2013) (CONTINUED)

N°	Short name	Name source	T-code	Series description
11	UMD	Kenneth French	1	Up Minus Down, sorted on momentum
12	Agric	Kenneth French	1	Agric industry portfolio
13	Food	Kenneth French	1	Agric industry portfolio
14	Beer	Kenneth French	1	Food industry portfolio
15	Smoke	Kenneth French	1	Smoke industry portfolio
16	Toys	Kenneth French	1	Toys industry portfolio
17	Fun	Kenneth French	1	Fun industry portfolio
18	Books	Kenneth French	1	Books industry portfolio
19	Hshld	Kenneth French	1	Household industry portfolio
20	Clths	Kenneth French	1	Clths industry portfolio
21	MedEq	Kenneth French	1	MedEq industry portfolio
22	Drugs	Kenneth French	1	Drugs industry portfolio
23	Chems	Kenneth French	1	Chems industry portfolio
24	Rubbr	Kenneth French	1	Rubber industry portfolio
25	Txtls	Kenneth French	1	Txtls industry portfolio
26	BldMt	Kenneth French	1	BldMt industry portfolio
27	Cnstr	Kenneth French	1	Construction industry portfolio
28	Steel	Kenneth French	1	Steel industry portfolio
39	Mach	Kenneth French	1	Mach industry portfolio
30	ElcEq	Kenneth French	1	ElcEq industry portfolio
31	Autos	Kenneth French	1	Autos industry portfolio
32	Aero	Kenneth French	1	Aero industry portfolio
33	Ships	Kenneth French	1	Ships industry portfolio
34	Mines	Kenneth French	1	Mines industry portfolio
35	oal	Kenneth French	1	Coal industry portfolio
36	Oil	Kenneth French	1	Oil industry portfolio
37	Util	Kenneth French	1	Util industry portfolio
38	Telcm	Kenneth French	1	Telcm industry portfolio
39	PerSv	Kenneth French	1	PerSv industry portfolio
40	BusSv	Kenneth French	1	BusSv industry portfolio
41	Hardw	Kenneth French	1	Hardw industry portfolio

FINANCIAL SERIES FROM JURADO, LUDVIGSON AND NG (2013) (CONTINUED)

N°	Short name	Name source	T-code	Series description
42	Chips	Kenneth French	1	Chips industry portfolio
43	LabEq	Kenneth French	1	LabEq industry portfolio
44	Paper	Kenneth French	1	Paper industry portfolio
45	Boxes	Kenneth French	1	Boxes industry portfolio
46	Trans	Kenneth French	1	Trans industry portfolio
47	Whlsl	Kenneth French	1	Whlsl industry portfolio
48	Rtail	Kenneth French	1	Rtail industry portfolio
49	Meals	Kenneth French	1	Meals industry portfolio
50	Banks	Kenneth French	1	Banks industry portfolio
51	Insur	Kenneth French	1	Insur industry portfolio
52	RlEst	Kenneth French	1	RlEst industry portfolio
53	Fin	Kenneth French	1	Fin industry portfolio
54	Other	Kenneth French	1	Other industry portfolio
55	1—2	Kenneth French	1	Portfolio sorted on (size, book-to-market)
56	1—4	Kenneth French	1	Portfolio sorted on (size, book-to-market)
57	1—5	Kenneth French	1	Portfolio sorted on (size, book-to-market)
58	1—6	Kenneth French	1	Portfolio sorted on (size, book-to-market)
59	1—7	Kenneth French	1	Portfolio sorted on (size, book-to-market)
60	1—8	Kenneth French	1	Portfolio sorted on (size, book-to-market)
61	1—9	Kenneth French	1	Portfolio sorted on (size, book-to-market)
62	1—high	Kenneth French	1	Portfolio sorted on (size, book-to-market)
63	2—low	Kenneth French	1	Portfolio sorted on (size, book-to-market)
64	2—2	Kenneth French	1	Portfolio sorted on (size, book-to-market)
65	2—3	Kenneth French	1	Portfolio sorted on (size, book-to-market)

FINANCIAL SERIES FROM JURADO, LUDVIGSON AND NG (2013) (CONTINUED)

N°	Short name	Name source	T-code	Series description
66	2—4	Kenneth French	1	Portfolio sorted on (size, book-to-market)
67	2—5	Kenneth French	1	Portfolio sorted on (size, book-to-market)
68	2—6	Kenneth French	1	Portfolio sorted on (size, book-to-market)
69	2—7	Kenneth French	1	Portfolio sorted on (size, book-to-market)
70	2—8	Kenneth French	1	Portfolio sorted on (size, book-to-market)
71	2—9	Kenneth French	1	Portfolio sorted on (size, book-to-market)
72	2—high	Kenneth French	1	Portfolio sorted on (size, book-to-market)
73	3—low	Kenneth French	1	Portfolio sorted on (size, book-to-market)
74	3—2	Kenneth French	1	Portfolio sorted on (size, book-to-market)
75	3—3	Kenneth French	1	Portfolio sorted on (size, book-to-market)
76	3—4	Kenneth French	1	Portfolio sorted on (size, book-to-market)
77	3—5	Kenneth French	1	Portfolio sorted on (size, book-to-market)
78	3—6	Kenneth French	1	Portfolio sorted on (size, book-to-market)
79	3—7	Kenneth French	1	Portfolio sorted on (size, book-to-market)
80	3—8	Kenneth French	1	Portfolio sorted on (size, book-to-market)
81	3—9	Kenneth French	1	Portfolio sorted on (size, book-to-market)
82	3—high	Kenneth French	1	Portfolio sorted on (size, book-to-market)
83	4—low	Kenneth French	1	Portfolio sorted on (size, book-to-market)

FINANCIAL SERIES FROM JURADO, LUDVIGSON AND NG (2013) (CONTINUED)

N°	Short name	Name source	T-code	Series description
84	4—2	Kenneth French	1	Portfolio sorted on (size, book-to-market)
85	4—3	Kenneth French	1	Portfolio sorted on (size, book-to-market)
86	4—4	Kenneth French	1	Portfolio sorted on (size, book-to-market)
87	4—5	Kenneth French	1	Portfolio sorted on (size, book-to-market)
88	4—6	Kenneth French	1	Portfolio sorted on (size, book-to-market)
89	4—7	Kenneth French	1	Portfolio sorted on (size, book-to-market)
90	4—8	Kenneth French	1	Portfolio sorted on (size, book-to-market)
91	4—9	Kenneth French	1	Portfolio sorted on (size, book-to-market)
92	4—high	Kenneth French	1	Portfolio sorted on (size, book-to-market)
93	5—low	Kenneth French	1	Portfolio sorted on (size, book-to-market)
94	5—2	Kenneth French	1	Portfolio sorted on (size, book-to-market)
95	5—3	Kenneth French	1	Portfolio sorted on (size, book-to-market)
96	5—4	Kenneth French	1	Portfolio sorted on (size, book-to-market)
97	5—5	Kenneth French	1	Portfolio sorted on (size, book-to-market)
98	5—6	Kenneth French	1	Portfolio sorted on (size, book-to-market)
99	5—7	Kenneth French	1	Portfolio sorted on (size, book-to-market)
100	5—8	Kenneth French	1	Portfolio sorted on (size, book-to-market)
101	5—9	Kenneth French	1	Portfolio sorted on (size, book-to-market)

FINANCIAL SERIES FROM JURADO, LUDVIGSON AND NG (2013) (CONTINUED)

N°	Short name	Name source	T-code	Series description
102	5—high	Kenneth French	1	Portfolio sorted on (size, book-to-market)
103	6—low	Kenneth French	1	Portfolio sorted on (size, book-to-market)
104	6—2	Kenneth French	1	Portfolio sorted on (size, book-to-market)
105	6—3	Kenneth French	1	Portfolio sorted on (size, book-to-market)
106	6—4	Kenneth French	1	Portfolio sorted on (size, book-to-market)
107	6—5	Kenneth French	1	Portfolio sorted on (size, book-to-market)
108	6—6	Kenneth French	1	Portfolio sorted on (size, book-to-market)
109	6—7	Kenneth French	1	Portfolio sorted on (size, book-to-market)
110	6—8	Kenneth French	1	Portfolio sorted on (size, book-to-market)
111	6—9	Kenneth French	1	Portfolio sorted on (size, book-to-market)
112	6—high	Kenneth French	1	Portfolio sorted on (size, book-to-market)
113	7—low	Kenneth French	1	Portfolio sorted on (size, book-to-market)
114	7—2	Kenneth French	1	Portfolio sorted on (size, book-to-market)
115	7—3	Kenneth French	1	Portfolio sorted on (size, book-to-market)
116	7—4	Kenneth French	1	Portfolio sorted on (size, book-to-market)
117	7—5	Kenneth French	1	Portfolio sorted on (size, book-to-market)
118	7—6	Kenneth French	1	Portfolio sorted on (size, book-to-market)
119	7—7	Kenneth French	1	Portfolio sorted on (size, book-to-market)

Financial Series from Jurado, Ludvigson and Ng (2013) (CONTINUED)

N°	Short name	Name source	T-code	Series description
120	7—8	Kenneth French	1	Portfolio sorted on (size, book-to-market)
121	7—9	Kenneth French	1	Portfolio sorted on (size, book-to-market)
122	8—low	Kenneth French	1	Portfolio sorted on (size, book-to-market)
123	8—2	Kenneth French	1	Portfolio sorted on (size, book-to-market)
124	8—3	Kenneth French	1	Portfolio sorted on (size, book-to-market)
125	8—4	Kenneth French	1	Portfolio sorted on (size, book-to-market)
126	8—5	Kenneth French	1	Portfolio sorted on (size, book-to-market)
127	8—6	Kenneth French	1	Portfolio sorted on (size, book-to-market)
128	8—7	Kenneth French	1	Portfolio sorted on (size, book-to-market)
129	8—8	Kenneth French	1	Portfolio sorted on (size, book-to-market)
130	8—9	Kenneth French	1	Portfolio sorted on (size, book-to-market)
131	8—high	Kenneth French	1	Portfolio sorted on (size, book-to-market)
132	9—low	Kenneth French	1	Portfolio sorted on (size, book-to-market)
133	9—2	Kenneth French	1	Portfolio sorted on (size, book-to-market)
134	9—3	Kenneth French	1	Portfolio sorted on (size, book-to-market)
135	9—4	Kenneth French	1	Portfolio sorted on (size, book-to-market)
136	9—5	Kenneth French	1	Portfolio sorted on (size, book-to-market)

FINANCIAL SERIES FROM JURADO, LUDVIGSON AND NG (2013) (CONTINUED)

N°	Short name	Name source	T-code	Series description
137	9—6	Kenneth French	1	Portfolio sorted on (size, book-to-market)
138	9—7	Kenneth French	1	Portfolio sorted on (size, book-to-market)
139	9—8	Kenneth French	1	Portfolio sorted on (size, book-to-market)
140	9—high	Kenneth French	1	Portfolio sorted on (size, book-to-market)
141	10—low	Kenneth French	1	Portfolio sorted on (size, book-to-market)
142	10—2	Kenneth French	1	Portfolio sorted on (size, book-to-market)
143	10—3	Kenneth French	1	Portfolio sorted on (size, book-to-market)
144	10—4	Kenneth French	1	Portfolio sorted on (size, book-to-market)
145	10—5	Kenneth French	1	Portfolio sorted on (size, book-to-market)
146	10—6	Kenneth French	1	Portfolio sorted on (size, book-to-market)
147	10—7	Kenneth French	1	Portfolio sorted on (size, book-to-market)

US MACROECONOMIC DATA FROM BOIVIN ET AL. (2013)

N°	Seriescode	T-Code	Series description
1	IPS10	5	Industrial production index - total index
2	IPS11	5	Industrial production index - products, total
3	IPS12	5	Industrial production index - consumer goods
4	IPS13	5	Industrial production index - durable consumer goods
5	IPS14	5	Industrial production index - automotive products
6	IPS18	5	Industrial production index - nondurable consumer goods
7	IPS25	5	Industrial production index - business equipment

US MACROECONOMIC DATA FROM BOIVIN ET AL. (2013) (CONTINUED)

N°	Seriescode	T-Code	Series description
8	IPS29	5	Industrial production index - defense and space equipment
9	IPS299	5	Industrial production index - final products
10	IPS306	5	Industrial production index - fuels
11	IPS32	5	Industrial production index - materials
12	IPS34	5	Industrial production index - durable goods materials
13	IPS38	5	Industrial production index - nondurable goods materials
14	IPS43	5	Industrial production index - manufacturing (sic)
15	PMP	1	Napm production index (percent)
16	PMI	1	Purchasing managers' index (sa)
17	UTL11	1	Capacity utilization - manufacturing (sic)
18	YPR	5	Pers income ch 2000 $, sa-us
19	YPDR	5	Disp pers income,billions of ch (2000) $, saar-us
20	YP@V00C	5	Pers income less trsf pmt ch 2000 $, sa-us
21	SAVPER	2	Pers saving,billions of $, saar-us
22	SAVPRATE	1	Pers saving as percentage of disp pers income, percent, saar-us
23	LHEL	5	Index of help-wanted advertising in newspapers (1967=100; sa)
24	LHELX	4	Employment: ratio; help-wanted ads: no. Unemployed clf
25	LHEM	5	Civilian labor force: employed, total (thous.,Sa)
26	LHNAG	5	Civilian labor force: employed, nonagric.Industries (thous., Sa)
27	LHTUR	1	Unemployment rate
28	LHU14	1	Unemploy.By duration: persons unempl.5 To 14 wks (thous., Sa)
39	LHU15	1	Unemploy.By duration: persons unempl.15 Wks + (thous., Sa)
30	LHU26	1	Unemploy.By duration: persons unempl.15 To 26 wks (thous., Sa)
31	LHU27	1	Unemploy.By duration: persons unempl.27 Wks + (thous, sa)
32	LHU5	1	Unemploy.By duration: persons unempl.Less than 5 wks (thous., Sa)
33	LHU680	1	Unemploy.By duration: average(mean)duration in weeks (sa)

US MACROECONOMIC DATA FROM BOIVIN ET AL. (2013) (CONTINUED)

N°	Series code	T-Code	Series description
34	LHUEM	5	Civilian labor force: unemployed, total (thous., Sa)
35	AHPCON	5	Avg hr earnings of prod wkrs: construction ($, sa)
36	AHPMF	5	Avg hr earnings of prod wkrs: manufacturing ($, sa)
37	PMEMP	1	Napm employment index (percent)
38	CES002	5	Employees on nonfarm payrolls - total private
39	CES003	5	Employees on nonfarm payrolls - goods-producing
40	CES004	5	Employees on nonfarm payrolls - natural resources and mining
41	CES011	5	Employees on nonfarm payrolls - construction
42	CES015	5	Employees on nonfarm payrolls - manufacturing
43	CES017	5	Employees on nonfarm payrolls - durable goods
44	CES033	5	Employees on nonfarm payrolls - nondurable goods
45	CES046	5	Employees on nonfarm payrolls - service-providing
46	CES048	5	Employees on nonfarm payrolls - trade, transportation, and utilities
47	CES049	5	Employees on nonfarm payrolls - wholesale trade
48	CES053	5	Employees on nonfarm payrolls - retail trade
49	CES088	5	Employees on nonfarm payrolls - financial activities
50	CES140	5	Employees on nonfarm payrolls - government
51	CES151	1	Average weekly hours of production or nonsupervisory workers on private
52	CES153	1	Nonfarm payrolls - goods-producing
53	CES154	1	Average weekly hours of production or nonsupervisory workers on private
54	CES155	1	Nonfarm payrolls - construction
55	CES156	1	Average weekly hours of production or nonsupervisory workers on private
56	CES275	5	Nonfarm payrolls - manufacturing
57	CES277	5	Average weekly hours of production or nonsupervisory workers on private
58	CES278	5	Nonfarm payrolls - manufacturing overtime hours
59	JQCR	5	Average weekly hours of production or nonsupervisory workers on private

US MACROECONOMIC DATA FROM BOIVIN ET AL. (2013) (CONTINUED)

N°	Series code	T-Code	Series description
60	JQCNR	5	Nonfarm payrolls - durable goods
61	JQCDR	5	Average hourly earnings of production or nonsupervisory workers on private
62	JQCSVR	5	Nonfarm payrolls - goods-producing
63	MOCMQ	5	Average hourly earnings of production or nonsupervisory workers on private
64	MSONDQ	5	Nonfarm payrolls - construction
65	PMDEL	1	Average hourly earnings of production or nonsupervisory workers on private
66	PMNO	1	Nonfarm payrolls - manufacturing
67	PMNV	1	Real personal cons exp quantity index (200=100), saar
68	HUSTSZ	4	Real personal cons exp-nondurable goods quantity index (200=100), saar
69	HSFR	4	Real personal cons exp-durable goods quantity index (200=100), saar
70	HSMW	4	Real personal cons exp-services quantity index (200=100), saar
71	HSNE	4	New orders (net) - consumer goods & materials, 1996 dollars (bci)
72	HSSOU	4	New orders, nondefense capital goods, in 1996 dollars (bci)
73	HSWST	4	Napm vendor deliveries index (percent)
74	EXRCAN	5	Napm new orders index (percent)
75	EXRUK	5	Napm inventories index (percent)
76	EXRUS	5	Housing starts: total new priv housing units (thous., Saar)
77	PMCP	1	Housing starts:nonfarm(1947-58);total farm&nonfarm(1959-) (thous., Sa)
78	PW561	5	Housing starts:midwest(thous.U.)S.A.
79	PWCMSA	5	Housing starts:northeast (thous.U.) S.A.
80	PWFCSA	5	Housing starts:south (thous.U.) S.A.
81	PWFSA	5	Housing starts:west (thous.U.) S.A.
82	PWIMSA	5	Foreign exchange rate: canada (canadian $ per u.s.$)
83	PUNEW	5	Foreign exchange rate: united kingdom (cents per pound)
84	PUS	5	United states;effective exchange rate(merm) (index no.)
85	PUXF	5	Napm commodity prices index (percent)

US MACROECONOMIC DATA FROM BOIVIN ET AL. (2013) (CONTINUED)

N°	**Series code**	**T-Code**	**Series description**
86	PUXHS	5	Producer price index: crude petroleum (82=100, nsa)
87	PUXM	5	Producer price index:crude materials (82=100, sa)
88	PUXX	5	Producer price index:finished consumer goods (82=100, sa)
89	PUC	5	Producer price index: finished goods (82=100, sa)
90	PUCD	5	Producer price index: intermed mat.Supplies & components(82=100, sa)
91	PU83	5	Cpi-u: all items (82-84=100, sa)
92	PU84	5	Cpi-u: services (82-84=100, sa)
93	PU85	5	Cpi-u: all items less food (82-84=100, sa)
94	FSDJ	5	Cpi-u: all items less shelter (82-84=100, sa)
95	FSDXP	1	Cpi-u: all items less midical care (82-84=100, sa)
96	FSPCOM	5	Cpi-u: all items less food and energy (82-84=100, sa)
97	FSPIN	5	Cpi-u: commodities (82-84=100, sa)
98	FSPXE	1	Cpi-u: durables (82-84=100, sa)
99	FM1	5	Cpi-u: apparel & upkeep (82-84=100, sa)
100	FM2	5	Cpi-u: transportation (82-84=100, sa)
101	CCINRV	5	Cpi-u: medical care (82-84=100, sa)
102	UOMO83	1	Common stock prices: dow jones industrial average
103	FYGM3	1	S&p's composite common stock: dividend yield (% per annum)
104	FYGM6	1	S&p's common stock price index: composite (1941-43=10)
105	FYGT1	1	S&p's common stock price index: industrials (1941-43=10)
106	FYGT10	1	S&p's composite common stock: price-earnings ratio (%, nsa)
107	FYGT20	1	Money stock: m1(curr,trav.Cks,dem dep,other ck'able dep) (bil$, sa)
108	FYGT3	1	Money stock:m2(m1+o'nite rps,euro$,g/p&b/d mmmfs&sav&sm time dep (bil$, sa)
109	FYGT5	1	Consumer credit outstanding – nonrevolving (g19)
110	FYPR	1	Composite indexes leading index component index of consumer expectations

US MACROECONOMIC DATA FROM BOIVIN ET AL. (2013) (CONTINUED)

N°	Series code	T-Code	Series description
111	FYAAAC	1	Units: 1966.1=100 Nsa, confboard and u.Mich.
112	FYAAAM	1	Interest rate: u.S.Treasury bills,sec mkt,3-mo. (% Per ann, nsa)
113	FYAC	1	Interest rate: u.S.Treasury bills,sec mkt,6-mo. (% Per ann,nsa)
114	FYAVG	1	Interest rate: u.S.Treasury const maturities,1-yr. (% Per ann, nsa)
115	FYBAAC	1	Interest rate: u.S.Treasury const maturities,10-yr. (% Per ann, nsa)
116	SFYGM3	1	Interest rate: u.S.Treasury const maturities, 20-yr. (% Per ann, nsa)
117	SFYGM6	1	Interest rate: u.S.Treasury const maturities, 3-yr. (% Per ann, nsa)
118	SFYGT1	1	Interest rate: u.S.Treasury const maturities, 5-yr. (% Per ann, nsa)
119	SFYGT5	1	Prime rate chg by banks on short-term business loans (% per ann, nsa)
120	SFYGT10	1	Bond yield: moody's aaa corporate (% per annum)
121	SFYAAAC	1	Bond yield: moody's aaa municipal (% per annum)
122	SFYBAAC	1	Bond yield: moody's a corporate (% per annum, nsa)
123	FYFF	1	Bond yield: moody's average corporate (% per annum)
124	Bspread10Y	1	Bond yield: moody's baa corporate (% per annum)

CANADIAN MACROECONOMIC DATA

N°	T-Code	Series description
1	6	CPI: all-items
2	6	CPI: all-items excluding eight of the most volatile components
3	6	CPI: all-items excluding food
4	6	CPI: all-items excluding energy
5	6	CPI: food and energy
6	6	CPI: energy
7	6	CPI: housing
8	6	CPI: goods

CANADIAN MACROECONOMIC DATA (CONTINUED)

N°	T-Code	Series description
9	6	CPI: durable goods
10	6	CPI: non-durable goods
11	6	CPI: Services
12	6	CPI: services excluding shelter services
13	5	Building permits: total residential and non-residential
14	5	Building permits: seasonally adjusted; Residential
15	5	Building permits: industrial
16	5	Building permits: commercial
17	4	Housing starts: total units
18	2	Average work week, manufacturing (Hours)
19	5	Housing index
20	5	New orders, durable goods
21	5	Retail trade, furniture and appliances
22	5	Shipment to inventory ratio, finished products
23	5	GDP at basic prices: all industries
24	5	GDP at basic prices: business sector industries
25	5	GDP at Basic Prices: non-business sector industries
26	5	GDP at basic prices: goods-producing industries
27	5	GDP at basic prices: service-producing industries
28	5	GDP at basic prices: industrial production
29	5	GDP at basic prices: durable manufacturing industries
30	5	GDP at basic prices: mining and oil and gas extraction
31	5	GDP at basic prices: construction
32	5	GDP at basic prices: manufacturing
33	5	GDP at basic prices: wholesale trade
34	5	GDP at basic prices: finance, insurance, real estate, rental and leasing
35	6	IPI: All manufacturing
36	6	IPI: total excluding food and beverage manufacturing
37	6	IPI: basic manufacturing industries
38	6	IPI: non-food (excluding basic manufacturing industries) manufacturing
39	6	IPI: primary metal manufacturing excluding precious metals
40	5	CommPI: total, all commodities

CANADIAN MACROECONOMIC DATA (CONTINUED)

N°	T-Code	Series description
41	5	CommPI: energy
42	5	CommPI: metals and Minerals
43	5	CommPI: forestry
44	5	toronto Stock Exchange, value of shares traded
45	5	Toronto Stock Exchange, volume of shares traded
46	5	Standard and Poor's/Toronto Stock exchange Composite Index, close
47	2	Toronto Stock Exchange, stock dividend yields (composite), closing quotations
48	5	FX: United States dollar, noon spot rate, average
49	5	FX: United States dollar, 30-day forward closing rate
50	5	FX: United States dollar, 180-day forward closing rate
51	5	FX: United States dollar, 1-year forward closing rate
52	5	FX: United Kingdom pound sterling, noon spot rate, average
53	5	FX: United Kingdom pound sterling, 90-day forward noon rate
54	5	FX: Swedish krona, noon spot rate, average
55	5	FX: Swiss franc, noon spot rate, average
56	5	FX: Japanese yen, noon spot rate, average
57	2	Bank rate
58	2	Forward premium or discount (-), United States dollar in Canada: 3 month
59	2	Prime corporate paper rate: 3 month
60	2	Government of Canada marketable bonds, average yield: 1-3 year
61	2	Government of Canada marketable bonds, average yield: 3-5 year
62	2	Government of Canada marketable bonds, average yield: 5-10 year
63	2	Government of Canada marketable bonds, average yield: over 10 years
64	2	Treasury bill auction–average yields: 3 month
65	2	Treasury bill auction–average yields: 6 month
66	2	Average residential mortgage lending rate: 5 year
67	5	Total, Canada s official international reserves
68	5	Convertible foreign currencies, United States dollars
69	6	Total business and household credit; seasonally adjusted
70	6	Household credit; seasonally adjusted

CANADIAN MACROECONOMIC DATA (CONTINUED)

N°	T-Code	Series description
71	6	Residential mortgage credit; seasonally adjusted
72	6	Consumer credit; seasonally adjusted
73	6	Business credit; seasonally adjusted
74	6	Short-term business credit; seasonally adjusted
75	6	Canadian dollar assets, total loans
76	6	Total personal loans
77	6	Business loans
78	6	M1B (gross)
79	6	Residential mortgages
80	6	M2+ (gross)
81	6	Chartered bank deposits, personal, term
82	6	Bankers acceptances
83	2	Unemployment rate (Rate); both sexes; 15 years and over
84	5	Total employed, all industries
85	5	EMP: Goods-producing sector
86	5	EMP: utilities
87	5	EMP: construction
88	5	EMP: manufacturing
89	5	EMP: services-producing sector
90	5	EMP: trade
91	5	EMP: transportation and warehousing
92	5	EMP: finance, insurance, real estate and leasing
93	5	EMP: professional, scientific and technical services
94	5	EMP: business, building and other support services
95	5	Imports, United States, including Puerto Rico and Virgin Islands
96	5	Imports, United Kingdom
97	5	Imports, European Union excluding the United Kingdom
98	5	Imports, Japan
99	5	Exports, United States
100	5	Exports, United Kingdom
101	5	Exports, European Union excluding the United Kingdom
102	5	Exports, Japan
103	5	Imports, total of all merchandise

Canadian Macroeconomic Data (CONTINUED)

N°	T-Code	Series description
104	5	Imports, sector 2 Energy products
105	5	Imports, sector 3 Forestry products
106	5	Imports, sector 4 Industrial goods and materials
107	5	Imports, sector 5 Machinery and equipment
108	5	Imports, sector 6 Automotive products
109	5	Exports, total of all merchandise
110	5	Exports, sector 2 Energy products
111	5	Exports, sector 3 Forestry products
112	5	Exports, sector 4 Industrial goods and materials
113	5	Exports, sector 5 Machinery and equipment
114	5	Exports, sector 6 Automotive products

REFERENCES

Ahn, S.C. and A.R. Horenstein (2013). "Eigenvalue Ratio Test for the Number of Factors". *Econometrica*, 81(3): 1203-1227.

Alessi, L., Barigozzi M., and M. Capasso (2010). "Improved Penalization for Determining the Number of Factors in Approximate Factor Models." *Statistics and Probability Letters*, 80: 1806-1813.

Amengual, D., and M. W. Watson (2007). "Consistent Estimation of the Number of Dynamic Factors in a Large N and T Panel." *Journal of Business and Economic Statistics*, 25: 91-96.

Bai, J., and S. Ng (2002). "Determining the Number of Factors in Approximate Factor Models." *Econometrica*, 70: 191-221.

Bai, J., and S. Ng (2007). "Determining the Number of Primitive Shocks in Factor Models." *Journal of Business and Economic Statistics*, 25(1): 52-60.

Bai, J., and S. Ng. (2013). "Principal Components Estimation and Identification of the Factors." *Journal of Econometrics*, 176(1): 18-29.

Barhoumi, K., O. Darné, and L. Ferrara (2013). "Testing the Number of Factors: An Assessment for a Forecasting Purpose." *Oxford Bulletin of Economics and Statistics*, 75: 64-79.

Bates, B. J., M. Plagborg-Moller, J. H. Stock, and M. W. Watson (2013). "Consistent Factor Estimation in Dynamic Factor Models with Structural Instability." *Journal of Econometrics*, 177(2): 289-304.

Bernanke, B., J. Boivin, and P. Eliasz (2005). "Measuring the Effects of Monetary Policy: A Factor-Augmented Vector Autoregressive (FAVAR) Approach." *Quarterly Journal of Economics*, 120: 387-422.

BOIVIN, J., and S. NG (2005). "Understanding and Comparing Factor Based Macroeconomic Forecasts." *International Journal of Central Banking*, 1(2): 117-152.

BOIVIN, J., and S. NG (2006). "Are More Data Always Better for Factor Analysis." *Journal of Econometrics*, 132: 169-194.

BOIVIN, J., M. GIANNONI, and D. STEVANOVIC (2013). "Dynamic Effects of Credit Shocks in a Data-Rich Environment." CIRANO Working Paper 2013s-11.

BREITUNG, J., and S. EICKMEIER (2011). "Testing for Structural Breaks in Dynamic Factor Models." *Journal of Econometrics,* 163(1): 71-84.

CHAMBERLAIN, G., and M. ROTHSCHILD (1983). "Arbitrage, Factor Structure, and Mean-variance Analysis on Large Asset Markets." *Econometrica*, 51: 1281-304.

CHEN, L., J. J. DOLADO, and J. GONZALO (2014). "Detecting Big Structural Breaks in Large Factor Models." *Journal of Econometrics*, 180(1): 30-48.

CHEN, N.-F., R. ROLL, and S. A. ROSS (1986). "Economic Forces and the Stock Market." *Journal of Business*, 59: 383-403.

CONNOR, G., and R. KORAJCZYK (1986). "Performance Measurement with the Arbitrage Pricing Theory: A New Framework for Analysis." *Journal of Financial Economics,* 15(3): 373-394.

CONNOR, G., and R. KORAJCZYK (1988). "Risk and Return in an Equilibrium APT: Application of a New Test Methodology." *Journal of Financial Economics,* 21(2): 255-289.

CONNOR, G., and R. A. KORAJZCYK (1993). "A Test for the Number of Factors in an Approximate Factor Model." *Journal of Finance*, 48: 1263-1291.

DUFOUR, J.-M., and D. STEVANOVIC (2013). "Factor-Augmented VARMA Models with Macroeconomic Applications." *Journal of Business & Economic Statistics*, 31(4): 491-506.

EICKMEIER, S., W. LEMKE, and M. MARCELLINO (2014). "A Classical Time Varying FAVAR Model: Estimation, Forecasting, and Structural Analysis." *Journal of the Royal Statistical Society*, Series A.

GUO-FITOUSSI, L., and O. DARNÉ (2014). "A Comparison of the Finite Sample Properties of Tion Rules of Factor Numbers in Large Datasets." LEMNA–University of Nantes Working Paper.

HALLIN, M., and R. LISKA (2007). "Determining the Number of Factors in the General Dynamic Factor Model." *Journal of the American Statistical Association*, 102: 603-617.

HANSEN, L.P., and S. F. RICHARD (1987). "The Role of Conditioning Information in Deducing Testable Restrictions Implied by Dynamic Asset Pricing Models." *Econometrica*, 55: 587-614.

JURADO, K., S. LUDVIGSON, and S. NG (2013). "Measuring Uncertainty." Forthcoming in *American Economic Review.*

KOROBILIS, D. (2013). "Assessing the Transmission of Monetary Policy Using Time-varying Parameter Dynamic Factor Models." *Oxford Bulletin of Economics and Statistics*, 75 (2).

Ludvigson, S., and S. Ng (2007). "The Empirical Risk-Return Relation: A Factor Analysis Approach." *Journal of Financial Economics*, 83: 171-222.

Ludvigson, S., and S. Ng (2009). "Macro Factors in Bond Risk Premia." *Review of Financial Studies*, 22 (12): 5027-5067.

Ludvigson, S., and S. Ng (2011). "A Factor Analysis of Bond Risk Premia." in A. Ulah and D. Giles Ed., *Handbook of Empirical Economics and Finance*, pp. 313-372.

Lutkepohl, H. (1984). "Linear Transformations of Vector ARMA Processes." *Journal of Econometrics*, 26: 283-293.

Onatski, A. (2009). "Testing Hypotheses About the Number of Factors in Large Factor Models." *Econometrica*, 77(5): 1447-1479.

Onatski, A. (2010). "Determining the Number of Factors from Empirical Distribution of Eigenvalues." *The Review of Economics and Statistics*, 92(4): 1004-1016.

Onatski, A. (2012). "Asymptotics of the Principal Components Estimator of Large Factor Models with Weakly Influential Factors." *Journal of Econometrics,* 168: 244-258.

Rao, C. R. (1955). "Estimation and Tests of Significance in Factor Analysis." *Psychometrika*, 20 (2).

Stock, J.H., and M.W. Watson (2002). "Forecasting Using Principal Components From a Large Number of Predictors." *Journal of the American Statistical Association*, 22 (12): 5027-5067.

Stock, J.H., and M.W. Watson (2005). "Implications of Dynamic Factor Models for VAR Analysis." NBER WP 11467.