



# UNSW Business School

## Working Paper

Never Stand Still

Business School

Economics

UNSW Business School Research Paper No. 2015 ECON 17

Changes in the Factor Structure of the U.S. Economy: Permanent Breaks or Business Cycle Regimes?

Luke Hartigan

This paper can be downloaded without charge from  
The Social Science Research Network Electronic Paper Collection:  
<http://ssrn.com/abstract=2642067>

# Changes in the Factor Structure of the U.S. Economy: Permanent Breaks or Business Cycle Regimes?

Luke Hartigan\*  
University of New South Wales

August 10, 2015

## Abstract

The factor structure of the U.S. economy appears to change over time. Unlike previous studies which suggest this is due to permanent structural breaks in factor loadings, I argue instead that the volatility and persistence of factor processes undergo recurring changes related to the business cycle. To capture this, I develop a two-step Markov-switching static factor estimation procedure and apply it to a well studied U.S. macroeconomic data set. I find strong support for Markov-switching in the factors processes, with switching variances being most dominant. Conditional on Markov-switching factor processes, tests for regime-dependent factor loadings show only moderate evidence of change. Overall, the results support regime-dependent factor processes as the main explanation for the diverging number of estimated factors in empirical applications and challenge the global linearity assumption implicit in large dimensional factor models of the U.S. economy.

**Keywords:** Approximate Factor Model; Large Data Sets; Markov Switching Model; Structural Breaks.

**JEL Classification:** C32; C38; C51; C55; E32.

---

\* *Correspondence:* Luke Hartigan, PhD Candidate, School of Economics, University of New South Wales, Sydney, 2052, NSW, Australia. Email: luke.hartigan@unsw.edu.au

# 1 Introduction

Factor modelling is a popular dimensional reduction technique that assumes a small number of latent variables (factors) explain most of the common variation observed in a panel of data. The degree of correlation between the factors and the variables in the panel is measured by the factor loadings. Together, they represent a factor structure, and changes in either the processes describing the factors or the loadings would result in a change in the panel's factor structure. Accordingly, any change in factor structure will mean that the estimated number of factors needed to adequately describe the panel of data will also change. A growing body of research has shown that permanent structural breaks in factor loadings can impact the number of estimated factors in a panel of data; however, this literature does not consider the possibility that processes describing the factors themselves could possibly be changing.

To be specific about changes in factor processes, if the volatility and persistence of the underlying variables in a panel of data are not constant, then this could result in changes in the volatility and persistence of the extracted factors themselves. Meanwhile, if a majority of underlying variables in the panel also show a similar change in volatility and persistence, then the factor loadings will not necessarily change. It is only if some portion of underlying variables display disproportionate changes in variances, persistence, and/or correlation, then this will likely result in a change in the factor loadings.

To allow for these alternative sources of change in factor structure, I develop a two-step Markov-switching approximate static factor estimation procedure, similar in spirit to the two-step method presented by Diebold and Rudebusch (1996) for (exact) dynamic factor models with Markov-switching intercepts. First, I extract the factors via principle components and then fit a series of univariate two-regime Markov-switching autoregressive models to each factor separately, treating the factors as data. Second, I compute regime-dependent factors by interacting the extracted factors with the estimated smoothed regime probabilities. Following this, I obtain regime-dependent factor loadings by regression. While this method is less efficient than a one-step method, it overcomes the 'curse of dimensionality' which can be an issue in estimating large dimensional Markov-switching models. I apply the procedure to the Stock and Watson (2005) data set and show that the factor structure of the U.S. economy appears to change with the peaks and trough of the business cycle and that the estimated number of static factors needed to adequately describe the U.S. economy has also changed over time. Taken together, these changes suggest that the stage of the business cycle can impact the degree of the common variation amongst a panel of macroeconomic time series.

A two regime model appears to capture the main time series properties of the estimated factors, with the first state corresponding to a low variability regime associated with expansionary periods, while the second state reflects a high variability regime, which is shown to match up with the standard NBER recession dates. Furthermore, although each regime-dependent factor is modelled separately, the estimated smoothed regime probabilities for each factor show a high degree of

similarity with each other and with the NBER recession dates.

Formal tests for Markov-switching in the processes describing the factors show strong support for nonlinearity, with switching variances having the largest impact. Meanwhile, tests for changes in factor loadings only indicate moderate support towards regime-dependent loadings. The results suggests that regime-dependent factor processes provide a better explanation than permanent structural breaks in factor loadings for the divergent number of estimated factors when using existing estimation methods. The findings also raises the question of whether the global linearity assumption implicit in static factor models is suitable in panels of data subject to instability such as macroeconomic time series.

The rest of this paper is organised as follows: Section 2 documents how the estimated number of static factors changes with peaks and troughs of the business cycle and how changing factor processes might be the cause of this. Section 3 presents a two-step Markov-switching estimation procedure for static factor models. Tests for Markov-switching and regime-dependent loadings as well as implications are discussed in Section 4, while Section 5 concludes.

## 2 Factor Models and the Business Cycle

For a single variable,  $x_{it}$ , in a panel of data  $X$  with dimensions  $i = 1, \dots, N$ ;  $t = 1, \dots, T$ , a factor model is defined as:

$$x_{it} = \Lambda_i F_t + \epsilon_{it} \tag{1}$$

In Equation (1) the factors are given by  $F_t$ , with dimension  $r \times 1$ , while  $\Lambda_i$ , with dimension  $1 \times r$ , are the factor loadings and measures the weight variable  $i$  puts onto each factor and is analogous to the slope coefficient in regression. The term  $\Lambda_i F_t$  is the common component of  $x_{it}$  whereas the residual term,  $\epsilon_{it}$ , represents the idiosyncratic or ‘unique’ component of  $x_{it}$ . Note, all of these parameters are *unobserved* and must be estimated jointly. The model is referred to as an approximate static factor model. The term ‘static’ refers to the contemporaneous relationship between the data  $x_{it}$  and the factors  $F_t$ ; however, note that  $F_t$  itself can be a dynamic process and this will be exploited later in Section 3. The term ‘approximate’ is used because in this set-up the idiosyncratic errors are permitted to be weakly correlated across  $i$  and  $t$  dimensions. This is in contrast to the traditional or ‘strict’ factor model approach which assumes the idiosyncratic terms are orthogonal to each other.

Estimating the number of static factors,  $r$ , is critical. The main estimation methods proposed in the literature use the fact that the first  $r$  eigenvalues of the covariance matrix of  $X$ , representing the ‘systematic’ part, are unbounded, whereas the remaining eigenvalues, representing the ‘non-systematic’ part, are bounded as both  $N, T \rightarrow \infty$ . Note, either  $X'X$  or  $XX'$  can be used in estimation since they share the same eigenvalues. Three main methods for determining  $r$  include: the  $IC_{P2}$  estimator (Bai and Ng 2002), the  $ED$  estimator (Onatski 2010), and the  $GR$  estimator

(Ahn and Horenstein 2013). All three estimators are based on estimating the number of factors using an integer sequence  $k$ , defined as  $0 \leq k \leq kmax$ , with  $kmax$  interpreted as the maximum number of factors to test for *a priori*. There is a fourth estimator, the *ABC* estimator of Alessi et al. (2010), which is based on the Bai and Ng (2002)  $IC_P$  estimator but requires two extra parameters to be determined related to tuning and stability checks. No guides exist for the correct specification of these additional parameters and so this estimator will not be considered here. We now explain each of the three main estimators in turn.

Bai and Ng (2002) suggested estimating the number of static factors,  $r$ , by minimising one of two model selection criterion functions ( $PC_P$  for Panel  $C_P$  and  $IC_P$  for panel information criterion). The two proposed estimators are linked to the eigenvalues of the covariance matrix of  $X$  with each estimator equal to the number of eigenvalues larger than a threshold value specified by a penalty function which depends on both  $N$  and  $T$ . The  $IC_{P2}$ , which has been shown to have good properties in simulations (Bai and Ng 2008), can be written as:

$$IC_{P2}(k) = \ln(\hat{\sigma}_k^2) + k \left( \frac{N+T}{NT} \right) \ln C_{NT}^2 \quad (2)$$

where  $\hat{\sigma}_k^2 = (NT)^{-1} \sum_{i=1}^N \sum_{t=1}^T (\hat{\epsilon}_{it}^k)^2$ ,  $C_{NT}^2 = \min[N, T]$ . The estimated number of factors,  $\hat{r}$ , is then given by the minimum  $IC_{P2}$  for the sequence  $k$ . Note, if true process describing the data is a dynamic factor model, then the Bai and Ng (2002) estimators will give an upper bound of the true number of factors in the panel as it will count both the factors and their lags.

Onatski (2010) proposed the ‘Edge Distribution’ (*ED*) estimator, which estimates the number of static factors using successive differences between eigenvalues of the covariance matrix of  $X$ . The threshold is found as the difference which is greater than some fixed value  $\delta > 0$  which needs to be calibrated. It can be written as:

$$\hat{r} = \max \{ i \leq k \max : \lambda_i - \lambda_{i+1} \geq \delta \} \quad (3)$$

where  $\lambda_i$  is the  $i$ th eigenvalue of the covariance matrix of  $X$ . The idea behind the test is that for any  $i > r$ , the difference in successive eigenvalues  $\lambda_i - \lambda_{i+1}$  converges to 0, while the difference  $\lambda_r - \lambda_{r+1}$  diverges to infinity. If the true DGP is a dynamic factor model then the *ED* should in principle capture the number of factors and their lags. Some advantages of the *ED* is that it works well in small samples when the amount of cross-sectional and temporal correlation in the idiosyncratic terms is relatively large. It also improves on the Bai and Ng (2002) estimators when the share of the observed variation in the panel attributed to the factors is small relative to the variation due to the idiosyncratic term (Onatski 2010). However, the *ED* may underestimate the number of factors if some of the factors are ‘weak’. Used in this context, weak is taken to mean that the smallest eigenvalue of the systematic part is close in magnitude to the largest eigenvalue of the non-systematic part.

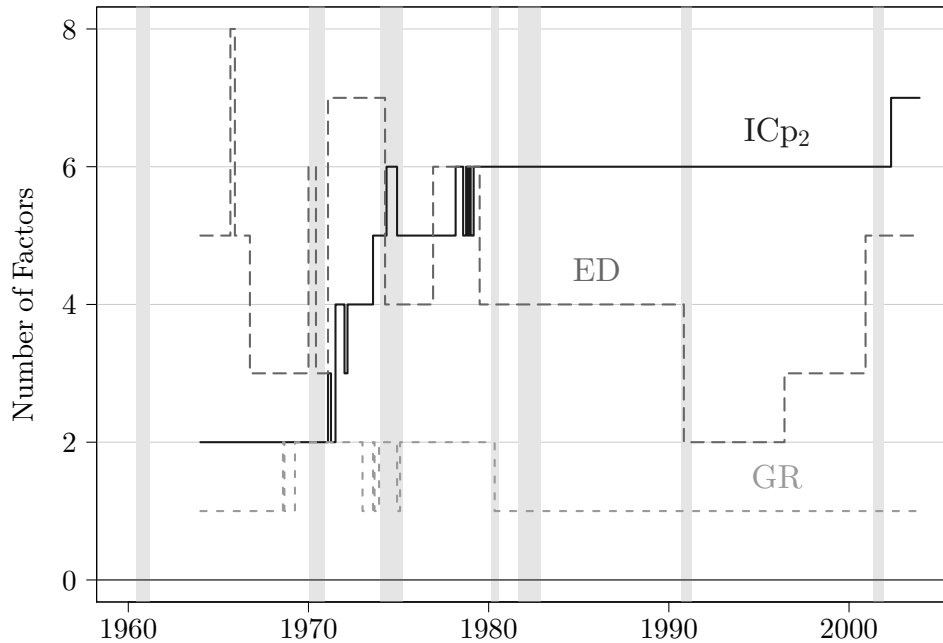
Ahn and Horenstein (2013) proposed two estimators: the ‘Eigenvalue Ratio’ ( $ER$ ) and the ‘Growth Ratio’ ( $GR$ ) and suggest estimating  $r$  as the maximiser of the ratio of two adjacent eigenvalues of the covariance matrix of  $X$  arranged in descending order. The authors report that the  $GR$  performs better than the  $ER$  in situations in which the panel has a dominant factor (in terms of the proportion of explained variation). This tends to be the case in empirical applications in macroeconomics so I focus solely on the  $GR$  estimator here. The  $GR$  can be written as:

$$GR(k) = \frac{\ln[V(k-1)/V(k)]}{\ln[V(k)/V(k+1)]} \quad (4)$$

where  $V(k) = \sum_{j=k+1}^{\min(N,T)} \lambda_j$ , the cumulative sum of the eigenvalues of the covariance matrix of  $X$ . The estimate of the number of factors,  $\hat{r}$ , is then given as the maximum  $GR$  for the sequence  $k$ . Note, the authors also suggest that these two estimators may be inappropriate in cases in which some factors have dynamic factor loadings of infinite order such as in Generalised Dynamic Factor models.

In their study Stock and Watson (2005) estimated seven static factors for the full sample of their data set (henceforth SW2005) using the  $IC_{P2}$  estimator. However, it is worth investigating if the estimated number of factors is sensitive to the sample period considered. To do this I follow the procedure first used by Bai and Ng (2007) and re-estimate the number of factors multiple times using all three estimators over an expanding time window starting at  $t = \text{January 1963}$  (an initial four year window), and expanding the length of the estimation window by one month until the end of the sample using  $kmax = 15$  at each step.

The results are presented in Figure 1. The starting point was chosen to maximise the length of time available for the sequential estimation procedure while also ensuring a relatively large initial sample size. Observe the  $IC_{P2}$  does not estimate the number of factors to be seven until towards the end of the time sample of the SW2005 data set. Another feature of Figure 1 is that the estimated number of factors over time from all three methods are quite different; nonetheless, one aspect common between all three estimators is the tendency for the estimated number of factors to change across peaks and troughs of the business cycle as judged by the NBER recession dates. This is especially true for the 1970s and early 1980s; a period renowned for its increased volatility as a consequence of the two oil shocks, rising unemployment and the significant movements in interest rates by the U.S. Federal Reserve as a way of mitigating historically high inflation at the time. One reason for the difference observed between the three estimators may relate to the some of the issues mentioned when discussing each one in turn, such as the the  $ED$ ’s tendency to underestimate  $r$  when some factors are weak and the  $GR$ ’s inability to deal with dynamics in factors.



NOTE: Shaded bars refer to NBER determined recession dates.

Figure 1: Recursive Estimated Number of Static Factors – SW2005 Data Set

It is worth asking why there are so many factors estimated in the panel of U.S. data, as Stock and Watson (2005) did in their original study, and what is causing the changes in the number estimated over time. Recently, the literature has come to understand that permanent structural breaks in factor loadings augment the factor space. The reason is that a factor model with breaks in loadings is equivalent to another factor model with an increased number of factors and constant factor loadings. The idea was first proposed by Breitung and Eickmeier (2011) and later proved by Chen et al. (2014). These authors, plus others such as Stock and Watson (2009), Corradi and Swanson (2014), Yamamoto and Tanaka (2014), Yamamoto (2014), and Han and Inoue (2015), have shown how breaks in factor loadings can inflate the estimated number of factors and each have proposed alternative tests for detecting a one-time common break in the factor loadings.

By focusing on structural breaks in factor loadings only, however, this literature downplays the idea that the processes describing the factors could also undergo change and, in addition, that this change may be recurrent and not necessarily permanent. This was first conjectured by Stock and Watson (2009) while Chen et al. (2014) discuss that their proposed test for structural breaks in factor models, which is based on the stability of regression coefficients of the first factor on the remaining factors, cannot differentiate between breaks in loadings, breaks in the covariance matrix of the factors or both. (The idea of breaks in factor dynamics was also mentioned by Han and Inoue 2015 in an earlier working paper version of their paper). Therefore, it might be the case that these proposed tests for breaks in loadings are just picking up the effects of a change in the factor

processes instead.

The factors themselves might be subject to change since the underlying macroeconomic data used in their estimation is well known to have instability (see the comprehensive studies of Stock and Watson 1996 for changes in the first moments and Sensier and van Dijk 2004 for changes in the second moments of many macroeconomic time series data frequently used in factor analysis). Because the factors represent the common variation in a panel of macroeconomic data, if the variances and/or covariances of the underlying variables change (due the alternation between recessions and expansions for instance), then these changes could cause a change in the time series properties of the extracted factors. For example, if the variances of the factors declines and this decline in variability is also reflected in a large proportion of the variables in the underlying panel, then the factor loadings will not change even though the factor processes have. Hence, an unexplored area of the research agenda in this literature relates to the possibility of changes in the factor processes themselves.

Chan et al. (2012) point out an insightful method for differentiating between the effects of a change in the loadings or in the factor processes. Provided there is only one source of instability in the factor model, they suggest comparing the estimated number of factors obtained from the full sample with those from different subsamples split at an estimated break point. If there is no change in factor loadings between the two points, then the number of factors should be consistently estimated for the full sample and each of the subsamples. Alternatively, provided the factors are stationary and there is a change in factor loadings, then the true number of factors can still be consistently estimated for each subsample; however, it will be overestimated for the full sample.

Hence, if the number of estimated factors in each subsample approximately sums to the number estimated for the full sample, that is,  $r \approx \sum_{j=1}^J r_j$  where  $j$  is the estimated number of break points, then it appears that changes in the factor loadings are the main cause of the divergent number of estimated factors, (which Han and Inoue 2015 label as a Type 1 Break). Alternatively, if  $r \approx r_j$  for  $j = 1, \dots, J$ , then it would appear that changes in the processes describing the factors are the reason for the different number of estimated factors, (which is called a Type 2 break by Han and Inoue 2015).

Furthermore, some support for this idea can also be found in the work of Bai and Ng (2007) where the authors' develop a test for the number of dynamic factors in a large  $N$ ,  $T$  setting. Using the SW2005 data set, the authors find that the number of estimated dynamic factors remains relatively stable over time at four, but that the estimated number of static factors diverges over time (similar to Figure 1). The divergence between the estimated number of static and dynamic factors could reflect a change in the persistence and/or variance of the factor processes over time. Note that a dynamic factor model with  $s$  lags on  $q$  dynamic factors can always be written as a static factor model with  $r = q(s + 1)$  static factors. The dimension of the static factors,  $r$ , is generally larger than that of the dynamic factors because it includes the leads and lags of the dynamic factors.



To investigate this idea in relation to the current study, I divide the SW2005 data set into five subsamples. Then, I re-estimate the number of static factors in each of these subsamples using the three different estimators to determine how  $r_j$  for  $j = 1, \dots, 5$  varies relative to the estimate of  $r$  over the combined samples inferred from Figure 1. To ascertain potential break points, I use the observed step increases in the  $IC_{P2}$  estimator from Figure 1 since Breitung and Eickmeier (2011) showed in simulation work that this estimator overestimates the true number of factors when there are breaks in factor loadings. There appears to be five break points to consider as subsamples: i) January 1960 to June 1971; ii) July 1971 to July 1973; iii) August 1973 to February 1978; iv) March 1978 to March 2002; and v) April 2002 to December 2003. The number of factors,  $r_j$ , are then estimated in each of these five subsamples using the three different methods with the results presented in Table 1:

Table 1: Estimated Number of Factors Across Subsamples

Subsample	Time Period	Sample Size	Estimator		
			$IC_{P2}$	$ED$	$GR$
1	1960M1–1971M6	138	2	7	1
2	1971M7–1973M7	25	2	7	1
3	1973M8–1978M2	55	5	4	2
4	1978M3–2002M3	289	6	4	1
5	2002M4–2003M12	21	1	5	2
All	1960M1–2003M12	528	7	5	1

*NOTE:* The subsamples were determined in reference to the step increases observed in the  $IC_{P2}$  estimator for the SW2005 data set in Figure 1. In each sub-sample  $kmax = 10$ .

There is a fair amount of variability in the estimated number of factors across the five subsamples. Focusing primarily on the  $IC_{P2}$  estimates, the first two subsamples show results consistent with what might be changes in factor loadings. This is because in each subsample, both  $r_1$  and  $r_2$  are estimated to be two, while for the two periods combined there is estimated to be four factors (Figure 1). However, the next two subperiods suggest otherwise, because the estimates for  $r_3$  and  $r_4$  rises to five and six respectively, when we would expect to see estimates closer to one in each subsample to mirror with the unit step rises in Figure 1. Collectively, the estimates for  $r_j$  in each of these two subsamples is identical to the combined sample estimates of five and six as would be the outcome if the change was predominately from changes in factor processes rather than changes in factor loadings. Additionally, the subsample March 1978 to March 2002 effectively encompasses the period referred to as the ‘Great Moderation’ which is known to correspond to a decline in the volatility of many macroeconomic series from around March 1984 (see Kim and Nelson 1999 and McConnell and Perez-Quiros 2000). This adds further weight to the notion that processes describing the factors might have changed during this period. Finally, the estimate of  $r_5$  for subsample five weighs in favour of a change in factor loading since the estimate of one would correspond with the unit step increase from six to seven observed in Figure 1 which occurred during that period.

As an additional check I also compare the factors estimated over the full sample to those from each of the subsamples using a fixed  $r$  in all cases. This idea is based on Theorem 3 (T3) in Stock and Watson (2002) and Proposition 1 (P1) in Chan et al. (2012). T3 suggests that even with time variation in the factor loadings, the subsample estimates of the static factors should nearly span the same factor space. Whereas P1 suggests that if the factor loadings change due to a big break, then factors associated to the changing loadings will be effectively zero before the break point, while these same factors will be non-zero after the break point. Hence, if the estimated subsample factors do not span the same spaces as the full sample estimated factors, then it suggests the full sample factors might be zero in that subsample and the change in the estimated number of factors is more likely due to a change in factor loadings. However, if the opposite is found, then it suggests changing factor processes are the cause.

One method to measure the spaces spanned by the factors in different subsamples is by using canonical correlation. Indeed, Stock and Watson (2009) do something similar, but they compare only two subsamples with the subjectively chosen break point of March 1984, corresponding the start of the Great Moderation. Canonical correlation analysis is used to study linear relationships between two sets of variables,  $x = (x_1, \dots, x_m)'$  and  $y = (y_1, \dots, y_n)'$ . If there are correlations between the two sets of variables, then canonical correlation analysis will find linear combinations of the  $x_i$  and  $y_j$  variables which have maximum correlation with each other (see Anderson 2003). We can define the sample squared canonical correlation coefficient,  $\hat{\rho}_j^2$ , such that  $1 > \hat{\rho}_1^2 \geq \hat{\rho}_2^2 \geq \dots \geq \hat{\rho}_p^2 > 0$  with  $p = \min[m, n]$ . The squared canonical correlations are just the characteristic roots of  $S_{xx}^{-1}S_{xy}S_{yy}^{-1}S'_{xy}$  where:  $S_{xx} = \text{var}(x)$ ,  $S_{yy} = \text{var}(y)$ , and  $S_{xy} = \text{cov}(x, y)$ . Table 2 present estimated sample squared canonical correlations for each subsample relative to the full sample of factors based on  $r = 7$  static factors. Squared canonical correlations close to one in value suggest that the subsample and the full sample of estimated static factors span nearly the same spaces whereas a value close to zero implies the subsample and full sample estimated factors do not span the same spaces.

Table 2: Subsample Squared Canonical Correlations

Factor	Subsample				
	1	2	3	4	5
1	0.998	0.993	0.998	0.999	0.996
2	0.989	0.992	0.995	0.996	0.986
3	0.978	0.980	0.989	0.994	0.957
4	0.960	0.865	0.945	0.991	0.886
5	0.902	0.660	0.877	0.980	0.774
6	0.871	0.454	0.728	0.973	0.389
7	0.003	0.169	0.037	0.727	0.236

*NOTE:* Each column entry represents the  $j$ th sample squared canonical correlation between the  $j$ th full sample factor and the  $j$ th subsample factor, for  $j = 1, \dots, 7$ .

There are clearly differences in the estimated sample squared canonical correlations across the five subsamples. The factors from the second and fifth subsamples appear to span *different* spaces than the full sample factors with squared canonical correlation estimates well below 0.9, although the relatively smaller sample sizes in subsamples two and five could be impacting these estimates. For the most part, the first five or six squared canonical correlations in subsamples one, three, and five appear relatively high (i.e., close to 0.9 and above), indicating that the factors in these subsamples seem to span the same spaces as the full sample factors. In comparison to the other subsamples, the factors estimated from the fourth subsample show the most similarity with the spaces spanned by the full sample factors, however, this too may be driven by its relatively larger sample size.

Note, with the exception of the seventh squared canonical correlation from subsample four, none of the other subsamples' seventh squared canonical correlation suggest that the last factor from each of those subsamples appear to span the same space as the seventh factor from the full sample. Indeed, the estimated seventh squared canonical correlation from the first subsample is very close to zero, suggesting that this factor could be a result of a change in factor loadings at some later point in time. Taken together, the results are not consistent with structural breaks in the factor loadings being the only reason for the divergent number of estimated factors in the SW2005 data set and gives some weight to the idea that maybe the processes describing the estimated factors change over time, especially for the 1970–1980 time period.

### **3 A Two-step Estimation Procedure For Markov Switching Factor Models**

Given the findings in Section 2, it is worthwhile trying to disentangle changes in factor loadings from those of factor processes, since the impact on the estimated number of static factors because of these alternative changes can be misinterpreted. If we had an indicator variable which could be used to distinguish between the two types of changes, then we could test each change separately. One way to achieve this is to treat the indicator as an unobserved variable and estimate it. If we assume it follows a first-order Markov process, then we can use a Markov-switching model for this purpose.

Hamilton (1989; 2005) showed that macroeconomic data can be successfully modelled with a switching mean and/or variance. Furthermore, Markov-switching dynamic factor models have a long history in the literature with the first use by Diebold and Rudebusch (1996) who proposed a two-step estimation method, conceptually similar to the one proposed here, with further contributions from Chauvet (1998) and Kim and Nelson (1998), but as yet there has been no attempt to use a Markov-switching framework within a static factor model. The idea of a switching static factor model is not unusual though and was suggested by Stock and Watson (2011) as an area for future research.

Recently Camacho et al. (2014) investigated the estimation properties of the two-step and one-step estimation methods in the dynamic factor model case and found that while the one-step method is generally preferred, the benefits diminish as the quality of the indicators increases (in the sense of a signal-to-noise ratio). Their findings suggest the increase in estimation error might not be an issue if we only want to get an inference of the regime probabilities as is the case here. Additionally, a two-step method overcomes the ‘curse of dimensionality’ that is an issue when trying to estimate large dimensional Markov-switching models.

### 3.1 Estimation Method

The method involves a two-step estimation process. In the first step, factors are extracted using principal components and univariate Markov-switching AR (MS-AR) models are estimated for each factor separately treating them as ‘data’ which is permissible as long as both  $N$  and  $T$  are large. The principal components estimators of the factors and loadings can be simultaneously estimated by solving the following minimisation problem:

$$\min_{\Lambda^k, F^k} V(k) = (NT)^{-1} \sum_{i=1}^N \sum_{t=1}^T \left( x_{it} - \Lambda_i^k F_t^k \right)^2 \quad (5)$$

The superscript  $k$  indicates the number of estimated factors and is not necessarily equal to the true number of factors  $r$ . There are two alternative but equal methods for estimating the factors in a panel using the method of principal components (for a detailed survey of factor models see Bai and Ng 2008). The first estimator of  $F$  is obtained by concentrating out  $\Lambda^k$  and imposing the normalisation  $F^{k'} F^k / T = I_k$  and that  $\Lambda^{k'} \Lambda$  is a diagonal matrix. The resulting estimator,  $\tilde{F}^k$ , is given by  $\sqrt{T}$  times the eigenvectors corresponding to the  $k$  largest eigenvalues of the  $T \times T$  matrix  $XX'$ . The corresponding factor loadings are given by  $\tilde{\Lambda}^{k'} = \tilde{F}^{k'} X / T$ .

The second method is obtained by first concentrating out  $F^k$  and then imposing the normalisation  $\Lambda^{k'} \Lambda^k / N = I_k$  and that  $F^{k'} F^k$  is a diagonal matrix. This alternative estimator is given by  $\bar{F}^k = X \bar{\Lambda}^k / N$ , with  $\bar{\Lambda}^k$  equal to  $\sqrt{N}$  times the eigenvectors corresponding to the  $k$  largest eigenvalues of the  $N \times N$  matrix  $X'X$ . Note, in both estimation methods the normalisation is necessary to achieve a unique identification for  $\Lambda^k$  and  $F^k$ . The first estimator is computationally efficient when  $T < N$ , while the second is efficient when  $N < T$ . Finally as in Bates et al. (2013), we can define the re-scaled factors  $\hat{F} = \bar{F} (\bar{F}' \bar{F} / T)^{1/2}$ . I will use this estimator of the factors (with the ‘hat’ notation removed) in the subsequent analysis.

I next assume the time series characteristics for each factor can be successfully captured by a Markov-switching AR(2) process where the persistence and the variance are allowed to switch between two regimes. The lag order for the AR was determined from examining *ACF* and *PACF* plots for each factor. Some factors, such as the first one, show a high level of persistence, while the lower-ordered factors show relatively less persistence. However, in the interests of parsimony

and to keep the model for each factor consistent, the same number of lags is used for each factor. Hence, the process describing each factor is defined as:

$$F_t = \phi_{1,S_t} F_{t-1} + \phi_{2,S_t} F_{t-2} + \eta_t, \quad \eta_t \stackrel{i.i.d.}{\sim} \mathcal{N}(0, \sigma_{S_t}^2) \quad (6)$$

This effectively says the estimated factors follow a mixture of Gaussian distributions with common location (mean) and different scale (variance) with the component weights determined by the regime probabilities and with a chance for the persistence of the process driving the factors to also change between regimes. Note, an intercept term not required since the factors have zero unconditional mean by construction. (However, as a robustness check, a Markov-switching Intercept model for each factor was investigated, but within each regime the constant term was found to be insignificant for all but the seventh factor. Additionally, the estimated value of the Likelihood for each process was less than the value estimated for the MS-AR model for each factor.) The discrete first order Markov process  $S_t$  is assumed to have two states: low and high variability ( $S_t = 1$  and  $S_t = 2$ ) respectively and is governed by the transition probabilities  $p_{ij} = Pr[S_t = j | S_{t-1} = i]$ . This measures the probability of moving from state  $i$  at time  $t - 1$  to state  $j$  at time  $t$ . The Gaussian Likelihood function for this process at time  $t$  conditional on  $S_t$  and the parameter vector  $\theta$  is then given by:

$$l_t(F_t | S_t; \theta) = \frac{1}{\sqrt{2\pi}\sigma_{S_t}} \exp \left[ \frac{-(F_t - \phi_{1,S_t} F_{t-1} - \phi_{2,S_t} F_{t-2})^2}{2\sigma_{S_t}^2} \right] \quad (7)$$

Following this, in the second step we calculate regime-dependent factors by interacting the extracted factors with the estimated smoothed probabilities,  $S_T$ , from the previous step which were computed using Kim (1994)'s smoothing algorithm based on the full sample. After which the regime-dependent factors are used to estimate regime-dependent loadings using regression. In particular, the estimate  $\hat{\Lambda}_i^{(j)}$  where  $j = \{1, 2\}$  can be calculated from the regression of  $\tilde{x}_{it}^{(j)}$  on  $\tilde{F}_t^{(j)}$ :

$$\hat{\Lambda}_i^{(j)} = \left[ \sum_{t=1}^T \begin{bmatrix} \tilde{F}_t^{(j)} \\ \tilde{F}_t^{(j)} \end{bmatrix} \begin{bmatrix} \tilde{F}_t^{(j)} \end{bmatrix}' \right]^{-1} \left[ \sum_{t=1}^T \begin{bmatrix} \tilde{F}_t^{(j)} \\ \tilde{F}_t^{(j)} \end{bmatrix} \begin{bmatrix} \tilde{x}_{it}^{(j)} \end{bmatrix} \right] \quad (8)$$

where:

$$\tilde{x}_{it}^{(j)} = x_{it}^{(j)} \times \sqrt{Pr[S_T = j | \hat{\theta}]} \quad (9)$$

and:

$$\tilde{F}_t^{(j)} = F_t \times \sqrt{Pr[S_T = j | \hat{\theta}]} \quad (10)$$

As reported in Hamilton (1994), Equation (8) describes  $\hat{\Lambda}_i^{(j)}$  as satisfying a weighted *OLS* orthogonality condition where each observation is weighted by the probability that it came from regime  $j$ . Note that in this regression there is one dependent variable but seven explanatory variables. To

compute a regime-dependent dependent variable as shown in Equation (9) I use  $\bar{Pr} [S_T = j | \hat{\theta}]$  which is equal to the average smoothed probability from all seven univariate Markov-switching AR(2) models. Other weighting schemes were investigated such as using the amount of variation in the data explained by each factor, but the results were very similar.

### 3.2 Data Set

I now turn to implementing this two-step estimation procedure with the SW2005 data set which is a well studied data set in the static factor model literature, see for example; Bai and Ng (2007), Breitung and Eickmeier (2011), Caner and Han (2014), and Yamamoto (2014). The SW2005 data set is comprised of monthly observations on 132 U.S. macroeconomic time series spanning the time period: January 1959–December 2003 and covering real (i.e., activity), nominal (i.e., prices) and financial categories.<sup>1</sup> Each series is transformed by removing outliers and taking logs and/or differencing so that the transformed series are approximately stationary (for precise details see Stock and Watson 2005). After transformations there are 528 observations for each series. Finally, the data are standardised to have zero unconditional mean and unit unconditional variance as is standard in factor analysis.

### 3.3 Empirical Results

Table 3 presents estimation results for the MS-AR(2) models for each factor and for the whole time period. The first thing to note is that the estimate for the standard deviation for each factor process is very different between the two regimes and are all highly significant in both regimes. Generally, the estimated sample standard deviation in the high variability regime ( $S_t = 2$ ) is more than twice that of the low variability regime ( $S_t = 1$ ) for factors one to four, and is very close to double for factors five to seven.

---

<sup>1</sup>The data set can be accessed from Mark Watson's web site: <http://www.princeton.edu/~mwatson/>.

Table 3: Markov-switching AR(2) Model Parameter Estimates

Parameter	Static Factor						
	1	2	3	4	5	6	7
$\phi_{1,(S_t=1)}$	0.484 (0.055)	0.493 (0.047)	0.347 (0.050)	-0.321 (0.056)	0.340 (0.058)	0.498 (0.052)	0.087 (0.081)
$\phi_{2,(S_t=1)}$	0.374 (0.054)	0.313 (0.047)	0.160 (0.049)	-0.070 (0.059)	0.313 (0.059)	0.338 (0.052)	-0.023 (0.096)
$\phi_{1,(S_t=2)}$	0.625 (0.089)	0.840 (0.090)	0.534 (0.146)	-0.310 (0.074)	0.312 (0.069)	0.389 (0.086)	0.279 (0.080)
$\phi_{2,(S_t=2)}$	0.174 (0.089)	-0.008 (0.080)	-0.009 (0.207)	-0.313 (0.075)	-0.166 (0.069)	0.319 (0.090)	0.028 (0.097)
$\sigma_{(S_t=1)}$	0.076 (< 0.001)	0.034 (< 0.001)	0.037 (< 0.001)	0.029 (< 0.001)	0.029 (< 0.001)	0.017 (< 0.001)	0.020 (< 0.001)
$\sigma_{(S_t=2)}$	0.162 (< 0.001)	0.071 (< 0.001)	0.079 (< 0.001)	0.067 (< 0.001)	0.047 (< 0.001)	0.033 (< 0.001)	0.034 (< 0.001)
$p_{11}$	0.97 (0.04)	0.99 (0.04)	0.99 (0.04)	0.97 (0.03)	0.98 (0.04)	0.95 (0.05)	0.94 (0.05)
$p_{12}$	0.03 (0.02)	0.01 (0.01)	0.01 (0.01)	0.03 (0.01)	0.02 (0.01)	0.05 (0.03)	0.06 (0.04)
$p_{21}$	0.08 (0.04)	0.04 (0.03)	0.06 (0.04)	0.04 (0.02)	0.02 (0.02)	0.11 (0.06)	0.09 (0.05)
$p_{22}$	0.92 (0.03)	0.96 (0.01)	0.94 (0.04)	0.96 (0.01)	0.98 (0.04)	0.89 (0.06)	0.91 (0.06)
$\mathcal{L}(\hat{\theta})$	473.33	945.62	914.45	924.59	989.57	1258.56	1176.64
$\text{Dur}_{(S_t=1)}$	35.75	112.84	90.30	38.71	63.41	21.78	16.18
$\text{Dur}_{(S_t=2)}$	12.15	23.54	15.64	22.54	45.33	8.70	11.29
$\text{Per}_{(S_t=1)}$	0.90	0.86	0.61	0.27	0.75	0.88	0.15
$\text{Per}_{(S_t=2)}$	0.83	0.83	0.52	0.56	0.41	0.79	0.36

*NOTE:* Estimation of the MS-AR(2) model for each factor was done using Maximum Likelihood with uniform initial values for the regime probability  $S_t$ . The regimes are labelled as, ‘Low Variance’ for  $S_t = 1$  and ‘High Variance’ for  $S_t = 2$ . Standard errors are in parentheses and were calculated using the Hessian matrix of the log-likelihood function.  $\mathcal{L}(\hat{\theta})$  refers to the estimated value of the Log-Likelihood. ‘Dur’ refers to the estimated average duration of each regime (in months) and is calculated as  $1/(1 - p_{ii})$  for  $i = 1, 2$ . ‘Per’ refers to the estimated persistence of each AR(2) process in each regime for each factor. It is computed as the absolute value of the largest eigenvalue of the  $2 \times 2$  companion matrix formed from the estimated AR(2) coefficients for each regime and each factor.

The autoregressive coefficients are all significant in the low variability regime for all factors except for the seventh factor. Interestingly, in the high variability regime the second autoregressive coefficient is insignificant for all factors while the first autoregressive coefficient is significant for factors one, two, three, and seven, and stands in contrast to the low variability regime estimation results. This suggests that the underlying process describing the factors in each regime may be different and potentially could be better captured using a time-varying dimension model (see Chan et al. 2012).

The duration of a regime, as measured by  $1/(1 - p_{ii})$   $i = \{1, 2\}$ , is estimated to be longer for the low variability regime compared to the high variability regime. Estimates for the duration of the

low variability regime range from 16 months on average for factor seven to 113 months on average for factor two. The persistence for each autoregressive process, as measured by the absolute value of the largest eigenvalue calculated from the  $2 \times 2$  companion matrix formed from the autoregressive coefficients, appears to also change between low and high variability regimes. Generally, persistence is higher in the low variability regime compared to the high variability regime. However, the reverse is true for factors four and seven. This is because the roots of the AR(2) process describing these two factors are complex and signifies that these two factor processes show *more* cyclical features in the high variability regime relative to the low variability regime. Plots of each smoothed probability and the high and low variability regime for each factor are presented in Figures 2–8.

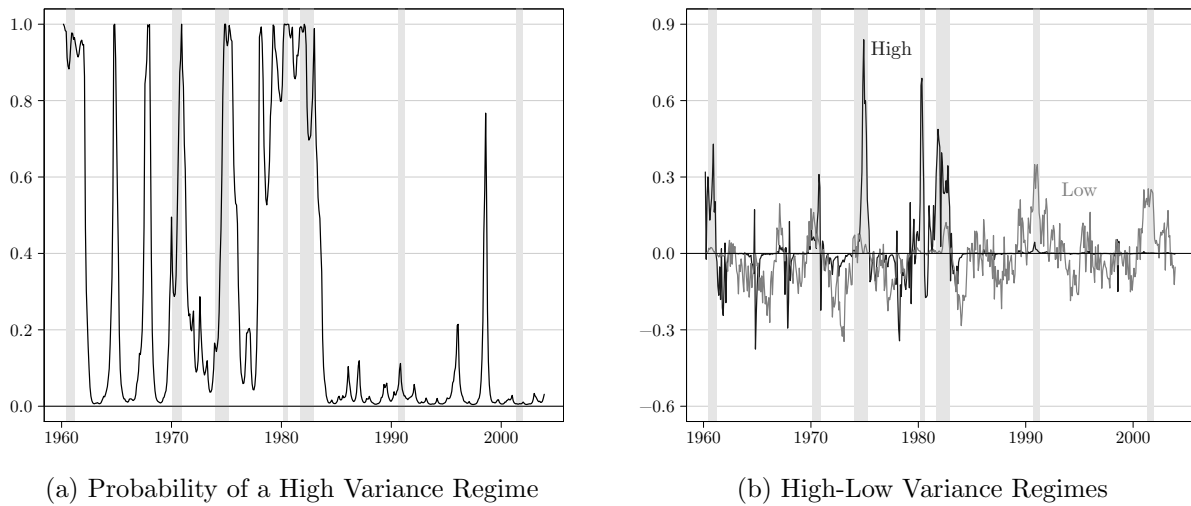


Figure 2: MS-AR(2) Probability and Regime-Dependent Factor 1

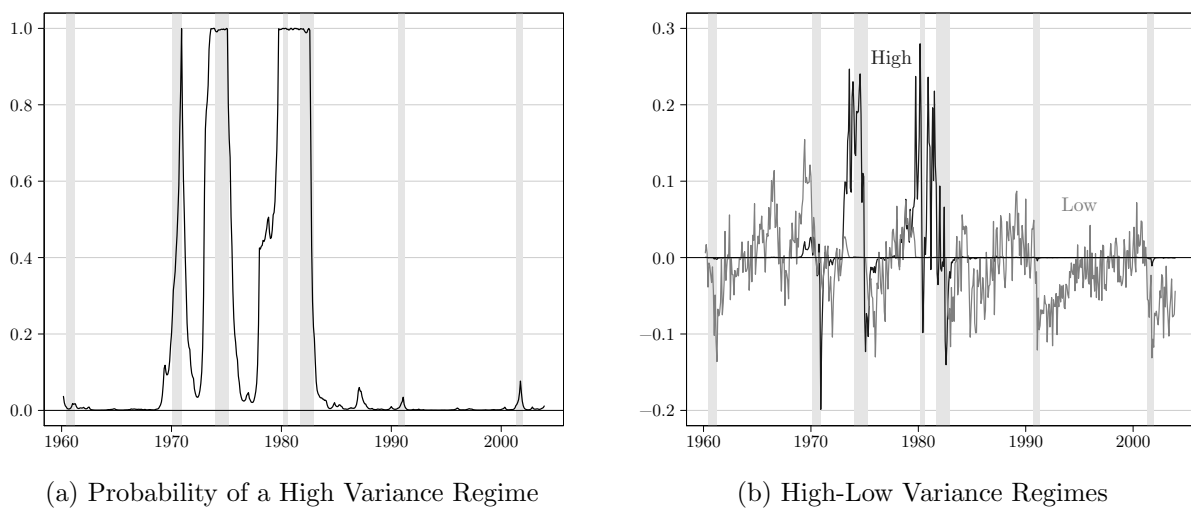
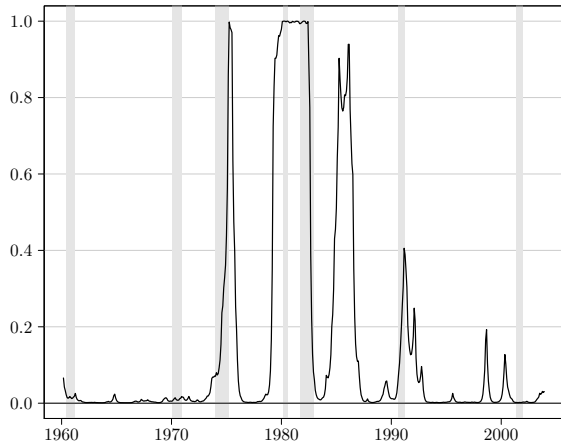
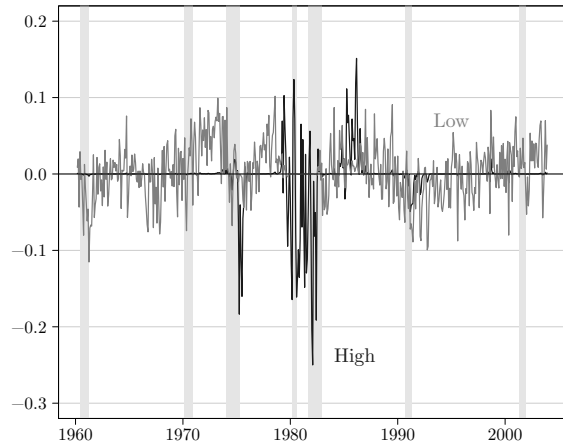


Figure 3: MS-AR(2) Probability and Regime-Dependent Factor 2



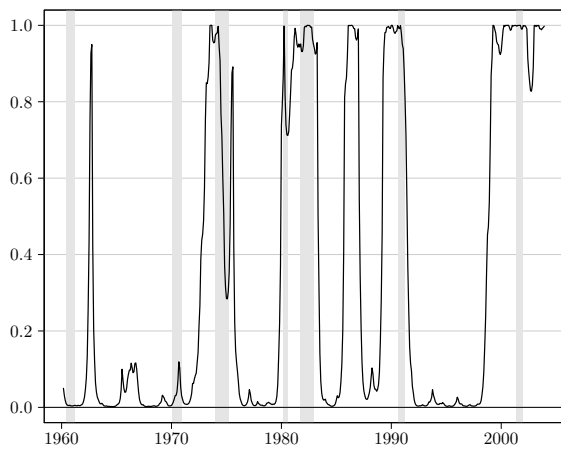


(a) Probability of a High Variance Regime

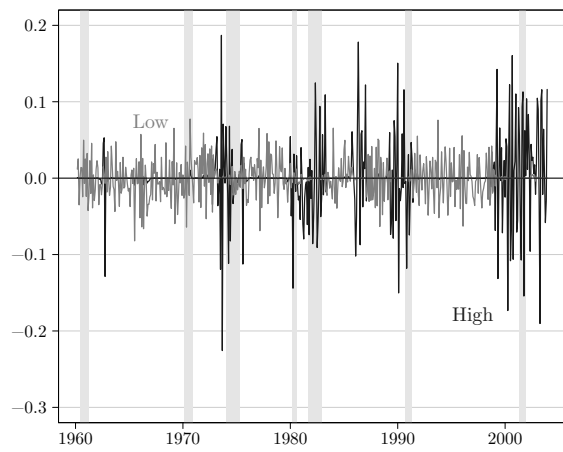


(b) High-Low Variance Regimes

Figure 4: MS-AR(2) Probability and Regime-Dependent Factor 3

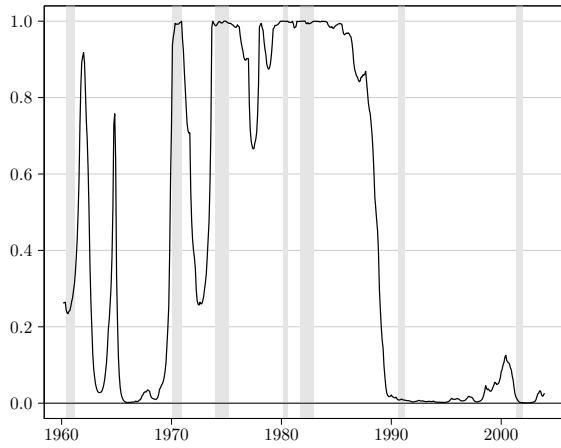


(a) Probability of a High Variance Regime

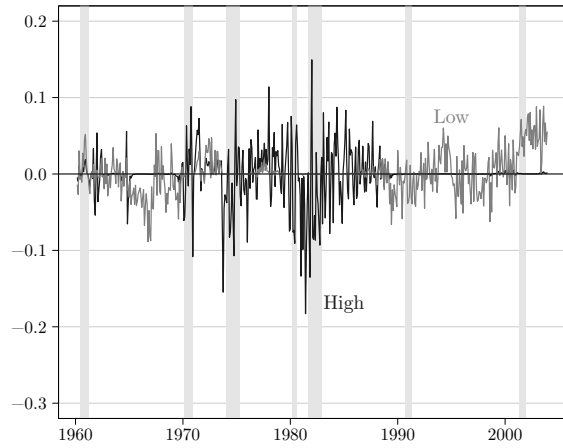


(b) High-Low Variance Regimes

Figure 5: MS-AR(2) Probability and Regime-Dependent Factor 4

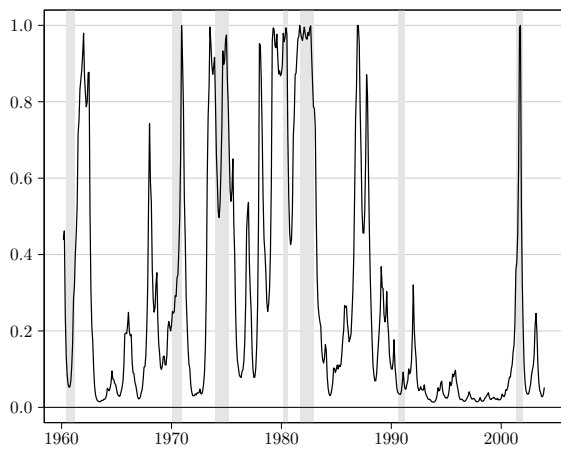


(a) Probability of a High Variance Regime

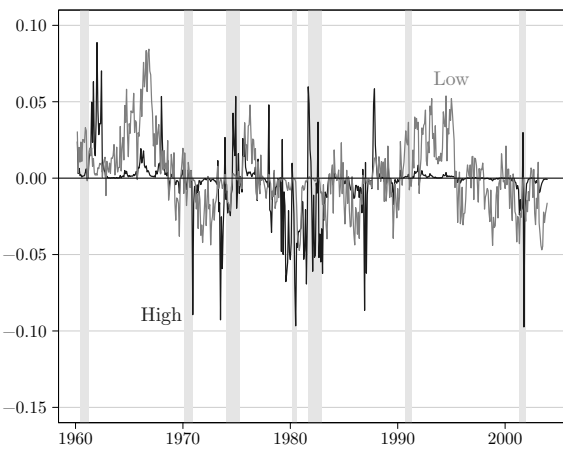


(b) High-Low Variance Regimes

Figure 6: MS-AR(2) Probability and Regime-Dependent Factor 5



(a) Probability of a High Variance Regime



(b) High-Low Variance Regimes

Figure 7: MS-AR(2) Probability and Regime-Dependent Factor 6

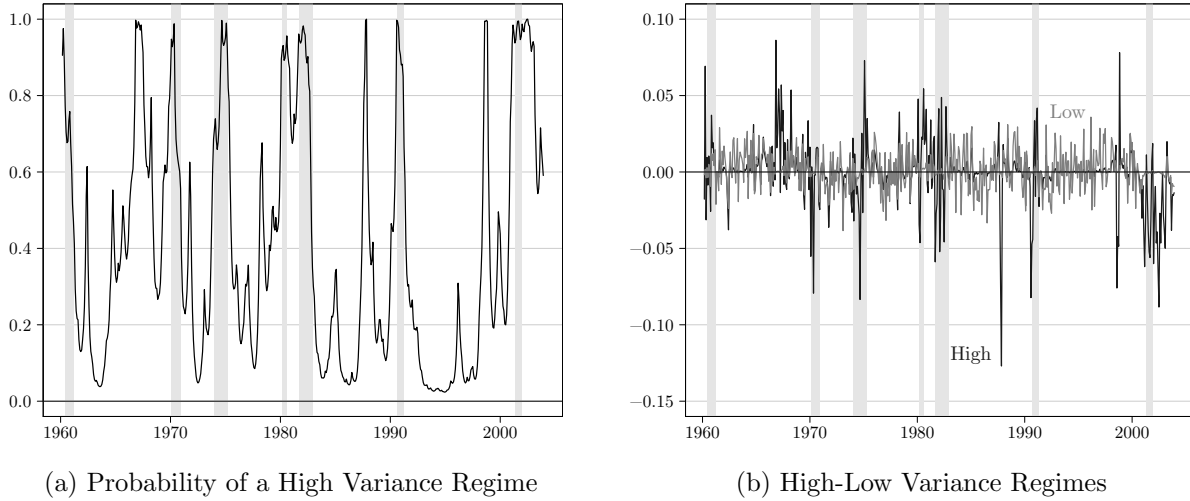


Figure 8: MS-AR(2) Probability and Regime-Dependent Factor 7

Generally the estimated smoothed regime probabilities capture the changes in the variability of the factors. The first factor, which appears mostly associated with real variables like Non-farm payrolls and industrial production, shows strong support for the Great Moderation. However, the first factor’s variance is shown to switch to the low variance regime from the end of 1982 which is two years before the acknowledged start of the Great Moderation in March 1984.

The second factor shows close linkages to variables representing interest rate spreads to the Federal Funds rate. The estimated smoothed probabilities for factor two show three main peaks, located in the early 1970s, mid 1970s and the early 1980s, which corresponds to the higher interest rate period associated with the time that Paul Volker was chairman of the Federal Reserve. Both the third and the fifth factors tend to describe changes in interest rate yields, and like factor two the estimated smoothed probabilities for both classify the 1970s/1980s period as being a high variance regime. The third Factor shows three prominent regime changes reminiscent of those observed for the second factor. In contrast, the fifth factor indicates that the whole 20 year period was characterised as a high variance regime.

The fourth factor appears to correspond to the variables related to changes in inflation (as measured by the CPI and the PCE price deflator) and shows frequent regime changes to the high variability regime at times of recessions. The last two factors, which seem to be linked to the housing sector (e.g., new home starts) and financial variables related to the stock market, show some estimation issues because the regime probabilities are not as well defined between the two regimes. This is evident because of the tendency for the estimated smoothed regime probability to hover in between the high and low variability regimes; however, the estimated smoothed probability for each factor does display a regime change at times when the U.S. economy has moved into recession.

Finally, all seven factors agree with the characterisation of the 1980–1982 period as being a high

variability regime as previously noted. This period is two years earlier than the beginning of the Great Moderation as well as the break date found by Breitung and Eickmeier (2011) using the same data set. However, the period of 1980–1982 more closely corresponds to the finding of a break point around 1979–1980 by Chen et al. (2014) who used a slightly different and more recent data set on the U.S. economy from Stock and Watson (2009). Furthermore, when treating the break point as an unknown nuisance parameter, Han and Inoue (2015) in an earlier working paper version and also using the Stock and Watson (2009) data set, find weak evidence towards a break point located around March 1984. Indeed, one of their sup-type tests ( $LM_{QS}$ ) suggest possible break points of September 1974 (at a 5% significance level) and June 1980 (at a 10% significance level).

Although the smoothed probability for each factor was estimated separately, there is a high degree of similarity between them and between each of them with the NBER determined recession dates. We can more formally measure their similarity using the concordance index proposed by Harding and Pagan (2002). This statistic, denoted by  $I_{jk}$ , quantifies the level of similarity as the fraction of time each smoothed probability ( $S_j$  where  $j = 1, \dots, 7$ ) and the NBER recession dates (the ‘reference’ series  $S_k$ ) remain in the same regime (Table 4). The index is bounded by the interval  $[0, 1]$ , with 0 indicating the two processes are exactly counter-cyclical and 1 indicating the two processes are exactly pro-cyclical and is defined as:

$$I_{jk} = T^{-1} \left[ \sum_{t=1}^T S_{jt} S_{kt} + \sum_{t=1}^T (1 - S_{jt}) (1 - S_{kt}) \right] \quad (11)$$

Table 4: Concordance Indices – Smoothed Regime Probabilities

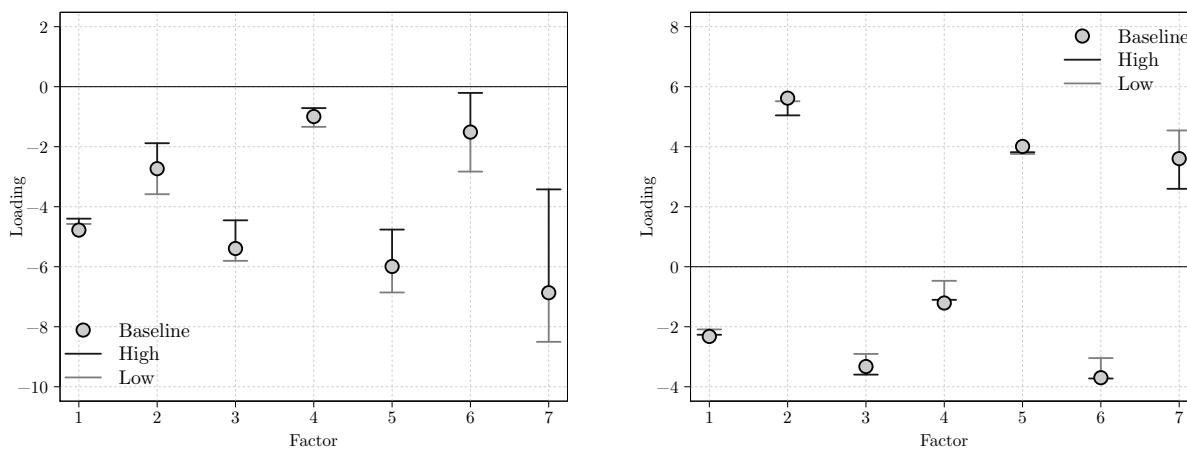
Statistic	Static Factor							Mean
	1	2	3	4	5	6	7	
$I_{jk}$	0.76	0.84	0.80	0.69	0.62	0.73	0.68	0.73
$\mathbb{E}[I_{jk}]$	0.66	0.73	0.75	0.61	0.55	0.65	0.56	0.65

*NOTE:* The Concordance index is a measure of the fraction of time the estimated smoothed probability for each factor ( $j = 1, \dots, 7$ ) and the NBER Recession dates ( $k$ ) are simultaneously in the same state of low variability ( $S_t = 1$ ) or high variability ( $S_t = 2$ ).

For comparison, the expected value of the index assuming the two processes were independent is also presented in the second row of Table 4. Each factor has a concordance index relatively close to 1 with the first three factors having a concordance index above 0.75 (with factor 2 being the highest) and the remaining factors at least above 0.6. It is also evident from Figure 3 and Figure 4 as well as Table 4 that the transitions between regimes for the second and third factors are very similar to each other.

Given the results of the Markov-switching models, it is useful to also compare how the factor loadings differ between the high and low variability regimes as well as the ‘baseline’ of no regime

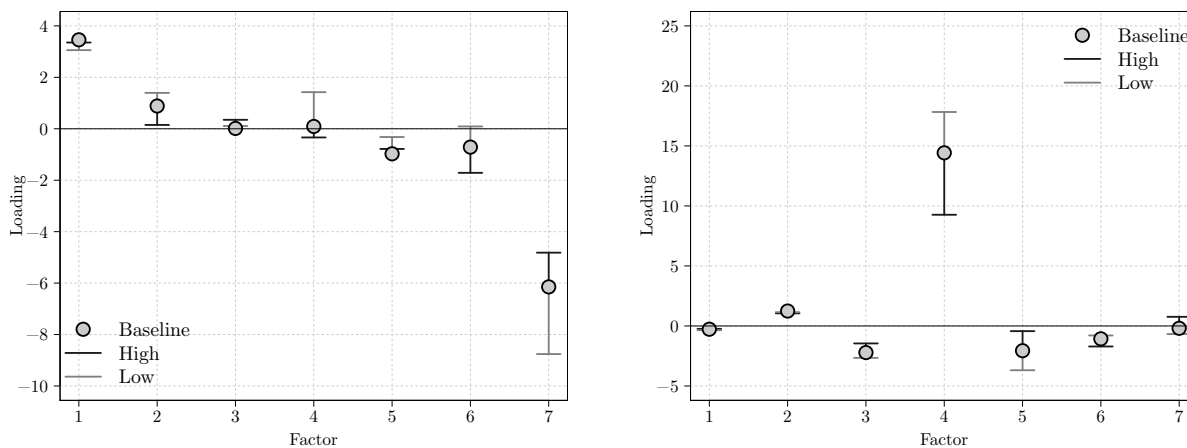
changes. This is illustrated in Figures 9–10 which plots the regime-dependent factor loadings for a selection of relatively important variables: industrial production (IPS10), the Federal Funds rate (FYFF), the unemployment rate (LHUR) and the PCE price deflator (GMDC). The dot-point for each loading in each plot represents the estimated value of each factor loading for these four variables assuming no regime switching. In comparison, the dark and light shaded bands highlight the estimated magnitude of the regime-dependent factor loadings for each of the four variables in the high and low variability regime respectively. The plots help illustrate the difference in the estimated regime-dependent factor loadings between the high and low-variability regimes and the baseline.



(a) Industrial Production – Total

(b) Federal Funds Rate

Figure 9: Regime-Dependent Factor Loadings – ISP10 & FYFF



(a) Unemployment Rate – All Workers

(b) PCE Price Deflator

Figure 10: Regime-Dependent Factor Loadings – LHUR & GMDC

What is notable is that for some of these variables, there does not appear to be much difference

in factor loadings between the two different regimes. This is especially true for the Federal Funds rate. Furthermore, a proportion of the factors appear unimportant in both regimes for some of these variables, such as the PCE price deflator, with estimated loadings close to zero in each regime. However, relatively large changes are observed for the fourth loading for the PCE price deflator and the seventh loading for industrial production. Finally, when there are divergences in the regime-dependent loadings, these tend to be related to the lower-order factors which is most evident with loadings five to seven for industrial production and the seventh loading for the unemployment rate.

We can use these results to understand the impact that changes in factor processes can have on factor loadings. If the factor processes change and the underlying observables also change by the same proportion, then we would not expect to observe any substantial changes in the estimates of the factor loadings in each regime but would instead observe regime switching in the factor processes. However, a change in factor loadings suggests that what is happening is more complicated than a simple rescaling of the variables and factors. Instead, it suggests the correlation structure of the observables is changing.

## 4 Testing for Regime Dependence in Static Factor Models

Given the preceding results from the two-step estimation method using the SW2005 data set, I now formally test for Markov-switching parameters in models of the factors and for changes in the factor loadings given the regime-dependent factors.

### 4.1 Do the Factor Processes Change over Time?

The previous results suggest the presence of Markov-switching, especially in the variance of each process. However, I test this formally in this section. Tests for Markov-switching are hampered by the presence of nuisance parameters which are only identified under the alternative hypothesis of Markov-switching as well as the additional issue of identically zero scores at the null hypothesis (see for example Hansen 1992). Nonetheless, the test introduced by Carrasco et al. (2014) (henceforth CHP) overcomes these problems and has previously been applied to macroeconomic data by Hamilton (2005) and Morley and Piger (2012) for example.

It is an information-matrix-based test for constancy of parameters in random coefficient models which also covers Markov-switching models. The test is based on functions of the first two derivatives of the likelihood evaluated under the null of no Markov-switching and the autocorrelations of the process describing the random parameters ( $\rho$ ). One advantage of the CHP test is that it only requires the estimation of the model under the null hypothesis. Furthermore, the authors' results indicate that the test has improved properties when the alternative allows for Markov-switching in

the variances of the process. The supremum form of the test statistic (supTS) is calculated as:

$$\text{supTS} = \sup \frac{1}{2} \left[ \max \left( 0, \frac{\Gamma_T^*}{\sqrt{\hat{\epsilon}^{*'} \hat{\epsilon}^*}} \right) \right] \quad (12)$$

Where  $\Gamma_T^* = \sum_{t=1}^T \mu_{2,t}^*(\beta, \theta) / \sqrt{T}$  and  $\hat{\epsilon}^*$  is the residual of the regression of  $\mu_{2,t}^*$  on the score of the likelihood function of the process under the null of no Markov-switching  $l_t^{(1)}(\hat{\theta})$ . We can calculate  $\mu_{2,t}^*(\beta, \theta)$  using:

$$\mu_{2,t}^*(\beta, \theta) = \frac{1}{2} h' \left[ \left( \frac{\partial^2 l_t}{\partial \theta \partial \theta'} + \left( \frac{\partial l_t}{\partial \theta} \right) \left( \frac{\partial l_t}{\partial \theta} \right)' \right) + 2 \sum_{s < t}^T \rho^{(t-s)} \left( \frac{\partial l_t}{\partial \theta} \right) \left( \frac{\partial l_s}{\partial \theta} \right)' \right] h \quad (13)$$

Here,  $l_t$  is the value of the likelihood at time  $t$  and  $h$  is a vector used to specify which parameters are allowed to change under the alternative. To test for switching variance I set  $h = (0, 0, 1)'$ , where 1 indicates that parameter ( $\sigma$ ) is permitted to change while the other two parameters ( $\phi_1$  and  $\phi_2$ ) are held fixed. To test for switching in the two autoregressive parameters as well as switching in all three parameters, I follow Carrasco et al. (2014) and generate the respective two or three elements in  $h$  uniformly over the unit sphere 100 times.

To use the supTS, we must also compute empirical critical values and  $p$ -values for the test statistic by parametric bootstrap. In doing so, I follow the strategy detailed in Carrasco et al. (2014). First, AR(2) models are fitted by Maximum Likelihood to each of the seven estimated factors under the null of no Markov-switching. From these estimates I compute the SupTS for each factor. Next, I generate  $S = 1000$  samples based on the Maximum Likelihood estimate for each AR(2) process with Gaussian innovations under the null of no Markov-switching for each factor. Then, for each bootstrapped sample, the Maximum Likelihood estimator for that simulated DGP is calculated and the  $\text{supTS}^{(s)}$   $s = 1, \dots, 1000$  is maximised numerically with respect to the nuisance parameter  $\rho \in [0.2, 0.8]$  to capture persistence, similar to Hamilton (2005). Finally, the empirical critical value for a nominal size  $\alpha$  was computed by finding the  $(1 - \alpha) \%$  quantile. The empirical  $p$ -value was then calculated as the proportion of simulated  $\text{supTS}^{(s)}$  which exceed the supTS for the actual process being examined. The results of the testing procedure are displayed in Table 5.

Table 5: CHP Test for Markov-switching Parameters

	Static Factor						
	1	2	3	4	5	6	7
Markov-switching Dynamics							
supTS	5.54	5.06	6.01	6.22	5.99	2.18	1.81
5% c.v.	2.15	2.24	2.20	2.33	2.19	2.13	2.41
<i>p</i> -value	0.00	0.00	0.00	0.00	0.00	0.05	0.09
Markov-switching Variance							
supTS	17.93	21.87	15.62	17.10	8.04	8.31	4.05
5% c.v.	1.14	1.32	1.28	1.24	1.27	1.14	1.33
<i>p</i> -value	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Markov-switching Dynamics and Variance							
supTS	18.52	22.15	15.97	23.08	9.16	8.55	5.39
5% c.v.	2.21	2.10	2.18	2.13	1.99	2.02	1.87
<i>p</i> -value	0.00	0.00	0.00	0.00	0.00	0.00	0.00

*NOTE:* The null hypothesis of the CHP test is no Markov-switching for each factor. The supTS for each factor and each test was calculated using  $\rho \in [0.2, 0.8]$ . The empirical 5% critical value for each test was calculated as the 95% quantile of the sample distribution of the supTS from 1000 replications of the AR(2) model under the null for each factor and each test. The empirical *p*-value for each test was calculated as the proportion supTS statistics greater than the observed test statistic.

For all factors the null hypothesis of no Markov-switching is strongly rejected for both cases of switching dynamics and variance and solely switching variances. Indeed, the largest empirical supTS<sup>(s)</sup> found in both tests was well below the calculated supTS for each factor. For the case of switching dynamics only, the evidence is less conclusive. While the first five factors reject the null, the sixth factor is borderline with a empirical *p*-value of 0.049 whereas the seventh factor does not reject the null at conventional levels of significance. Furthermore, the test statistics for the test of only switching dynamics are much smaller in magnitude than those for the other two tests. These findings suggest that switching variance is the most dominant changing parameter for each factor process, with some support for switching dynamics for the first five factors at least.

The test results also suggest some delineation between the first four factors to the remaining three factors from the full-sample of the SW2005 data set. Indeed, the supTS value for seventh factor is the smallest in each of the three different tests, and while it is still significant, it does indicate that the evidence against the null is less overwhelming than it is for the second or fourth factors for instance. When viewed together with the previous results from the canonical correlations analysis and Figures 9–10, these findings indicate that a possible conclusion could be that the first four factors are more likely to support the notion of changes in the processes describing the factors, whereas the remaining three factors could in fact be a result of breaks in factor loadings.



## 4.2 Are Factor Loadings Regime-Dependent?

It was previously shown that there are some observed differences in factor loadings between the high and low variability regimes for some variables (e.g., industrial production), whereas others are less supportive (e.g., the Federal Funds rate). As in the previous Section, I now propose three methods to test for changes in the factor loadings given the processes describing the factors support Markov-switching parameters more formally. These include, a Wald test,  $W_i$ , a Likelihood Ratio test,  $LR_i$ , and a Lagrange Multiplier test,  $LM_i$ , for each variable  $i = 1, \dots, N$  in the SW2005 data set. I now define each in turn. The Wald test for changing loadings can be written as:

$$W_i = T \left( \hat{\Lambda}_i^{(1)} - \hat{\Lambda}_i^{(2)} \right)' \left( \hat{\Sigma}_i^{(1)} + \hat{\Sigma}_i^{(2)} \right)^{-1} \left( \hat{\Lambda}_i^{(1)} - \hat{\Lambda}_i^{(2)} \right) \sim \chi_{(r)}^2 \quad (14)$$

Where  $\hat{\Lambda}_i^{(j)}$   $j = \{1, 2\}$  is calculated via regression as described by Equation (8) and  $\hat{\Sigma}_i^{(j)}$   $j = \{1, 2\}$  is the estimated (HAC corrected) covariance matrix for the estimated regime-dependent factor loadings from the two different regimes, thereby allowing for the estimated covariance matrix to be different in the two regimes.

The Likelihood Ratio test statistic is formed by first estimating a regression using the non-regime switching form of the factor model as in Equation (1) and labelling this the ‘restricted model’ and then obtain the value of the estimated likelihood for this model,  $\hat{\mathcal{L}}_R^i$ . Next, another regression is estimated for each variable incorporating the regime-dependent factors as calculated in Equation (10) which I label as the ‘unrestricted model’ and then obtain the value of the estimated likelihood from this second regression,  $\hat{\mathcal{L}}_{UR}^i$ . Finally, I compute the Likelihood Ratio test statistic as:

$$LR_i = -2 \left( \hat{\mathcal{L}}_R^i - \hat{\mathcal{L}}_{UR}^i \right) \sim \chi_{(r)}^2 \quad (15)$$

The Lagrange Multiplier test statistic can be constructed by taking the estimates of the idiosyncratic term from the ‘restricted’ regression model formed when computing the Likelihood Ratio test and then regressing these idiosyncratic terms on the regime-dependent factors calculated as in Equation (10) for all the variables in the SW2005 data set and then calculate the Lagrange Multiplier test statistic as:

$$LM_i = TR_i^2 \sim \chi_{(r)}^2 \quad (16)$$

Where  $T$  is the sample size and  $R_i^2$  is the  $R^2$  calculated from the  $i$ th regression of the idiosyncratic term on the regime-dependent factors for the  $i$ th variable. The collected results for all the variables in the SW2005 data set as well as for sub groupings using the significance levels  $\alpha = 5\%$  and the relatively more conservative  $\alpha = 1\%$  are presented in Table 6. The classification of the sub groups follows those used by Yamamoto (2014).

There is a small difference between the Wald test results and those of the other two tests. This could be due to the Wald test explicitly accounting for different covariance matrices in each of

the two regimes which is not the case with the Likelihood Ratio or Lagrange Multiplier tests. At a nominal size of 5%, just under half of the SW2005 data set supports regime-dependent factor loadings. Because I am aggregating a large number of individual tests together, it makes sense to use a more conservative significance level for each individual test. At the 1% level, the proportion of rejections for the three tests declines to around one-third of the panel.

Table 6: Tests for Regime-Dependent Factor Loadings

Category	Count	$\alpha = 1\%$			$\alpha = 5\%$		
		<i>W</i>	<i>LR</i>	<i>LM</i>	<i>W</i>	<i>LR</i>	<i>LM</i>
Income/Consumption/Employment	39	0.26	0.18	0.18	0.38	0.44	0.41
Production/New orders/Inventories	25	0.44	0.32	0.32	0.60	0.48	0.48
Housing	10	0.60	0.70	0.70	0.70	0.80	0.80
Money/Credit	11	0.00	0.18	0.18	0.09	0.27	0.27
Financial	26	0.58	0.54	0.54	0.62	0.77	0.77
Prices	21	0.29	0.14	0.14	0.43	0.24	0.24
All sectors	132	0.36	0.31	0.31	0.48	0.49	0.48

*NOTE:* The numbers represent the proportion of variables for each group which reject the null hypothesis of constant factor loadings. The *Wald* statistics for  $\alpha = 1\%$  and  $\alpha = 5\%$  are based on HAC corrected covariance matrices using the Bartlett kernel.

When compared to previous results these proportions are slightly lower than those reported by Breitung and Eickmeier (2011) and Yamamoto (2014), who both used the same data set but only focused on testing for a one-time change in factor loadings, ignoring regime changing factor processes. For instance, when using their preferred *LM* and *sup-LM* tests and  $r = 7$  factors, Breitung and Eickmeier (2011) find 52% and 67% of the (outlier adjusted) SW2005 data set support changing factor loadings. In addition, Yamamoto (2014) finds that around 65% of the SW2005 data set have unstable factor loadings when using his preferred testing procedure (the ‘MBE-mod’ method).

In contrast, the results from the three proposed tests are in line with the findings of Stock and Watson (2009) who use a more recent data set and a Chow-test with the pre-determined break point of March 1984. At the 5% level, they report 41% of their panel reject the null of constant factor loadings while at the 1% level they find only 23% of their panel reject the null. What is notable from this is that Stock and Watson (2009)’s results are based on using only four factors, which suggests that it could be the case that the extra factors used by Breitung and Eickmeier (2011) might be causing this divergence between these sets of results.

Looking at the sub-categories, the majority of rejections of the null of constant factor loadings is from the three groups: ‘Income/Consumption/Employment’, ‘Production/New orders/Inventories’, and ‘Financial’ which incorporates stock market, interest rate and exchange rate time series. However, as a proportion of each sub-group total, the ‘Housing’ sub-group displays the largest percentage of series supporting regime-dependent factor loadings for all three tests followed by the

‘Financial’ sub group, which is similar to the findings of Stock and Watson (2009) and Yamamoto (2014) who each find variables from these two sub groups were more likely to have unstable factor loadings than other sub groups. In contrast, the ‘Money/Credit’ sub-category shows the smallest proportion of rejections in all three tests which is similar to the findings of Stock and Watson (2009) for this category.

Linking these results back to Figures 9–10; both industrial production and the PCE price deflator reject the null of constant factor loadings at 1% and 5% significance levels for all three tests, whereas the Federal Funds rate does not. The results for the unemployment rate series differs between the Wald test, which does not reject the null at both 1% and 5% levels, and the other two which do reject the null but only at the 5% level. With these results there is also some disagreement with previous studies, for example, Stock and Watson (2009) find little evidence for changing factor loadings for the Federal Funds rate whereas Breitung and Eickmeier (2011) reject the null of constant factor loadings for this same variable.

In summary, once we account for regime-dependent factor processes, the proposed three test results provide moderate evidence supporting regime changes for the factor loadings. Hence the conclusions of previous authors who focus exclusively on testing for a change in factor loadings could also be influenced by the presence of switching factor processes as well.

### 4.3 What are the Implications?

There is strong evidence that the dynamics and the variability of the estimated factors change with the peaks and troughs of the business cycle. Switching variances appear to be the main feature of change in the factor processes. High variance regimes are associated with recession periods, while the low variance regime associated with expansionary periods. There is also some evidence to suggest the persistence of the factors also changes between recessions (less persistent) and expansions (more persistent). On the other hand, there is only moderate evidence to support changes in factor loadings. One reason for this result could be that the factors are just describing the common features of the data. If a majority of the variables used to extract the factors also show the same changing variance or persistence characteristics, then only a minority of variables that do not share those *common* features will be likely to support changes in the factor loadings as judged by my proposed tests.

Another implication is that it appears, at least in reference to the SW2005 data set on the U.S. economy, that the observed increase in the estimated number of static factors does not simply reflect either changes in loadings or changes in factor processes, but reflects a combination of both, and these changes can take place at different times.

Finally, the results also question whether the global linearity assumption implicit in static factor models is suitable in panels of data subject to instability such as macroeconomic time series.

Perhaps instead what might be more appropriate is a ‘global’ non-linear process comprised of a ‘locally’ linear set of factor models (with different factor loadings and/or factors or even a different number of factors in each subsample) similar in concept to the mixtures of factor analyzers popular in the statistics literature (see for example McLachlan and Peel 2000). This might provide a new area of research for the factor model literature.

## 5 Conclusion

Factor models are a popular method for summarising the common variation contained in a large panel of macroeconomic data. However, the alteration of peaks and troughs of the business cycle can impact this common variation which results in a change in the factor structure describing the panel. The change in factor structure can come from two sources: changes in the factor loadings or changes in the factor processes. Both have consequences when trying to estimate the number of factors in the panel. Previous studies have attributed changes in factor loadings as a primary reason for this occurrence; however, these studies have not considered the possibility that the factors processes themselves change as the economy transitions between periods of expansion and contraction. Using data on the U.S. economy, I highlight that there are changes in the estimated number of factors between different time periods and across three established estimators. The changes correspond with peaks and troughs of the business cycle as measured by NBER recession dates. By comparing subsamples with full sample estimates, I show that some subsamples appear to support changes in factor loadings while other subsamples seem to support changes in the actual factors processes.

To help disentangle these two effects I proposed a two-step estimation procedure allowing for Markov-switching factor processes. In the first step, the factors are estimated using principal components and then individual MS-AR(2) processes for each factor are estimated treating the factors as data. In the second step, the estimated smoothed transition probabilities are used as an indicator variable to identify regime-dependent factors. These estimates are then employed to compute regime-dependent factor loadings using regression.

Tests for Markov-switching parameters in the factor processes displayed strong support, with switching variances the most important feature. Conditional on the extracted factors supporting regime switching, tests for regime-dependent factor loadings were proposed. The results showed moderate support against constant factor loadings, with around one-third of the SW2005 data set supporting regime-dependent factor loadings.

Finally, the assumption of a globally linear factor model is challenged, and alternative models which allow some non-linearity as proposed in this paper are recommended as avenues for future research.

## Acknowledgements

I am grateful to James Morley, Timothy Neal, and Chris Heaton, for helpful comments and suggestions. All errors remain my own.

## References

- Ahn, S. C. and A. R. Horenstein (2013). Eigenvalue ratio test for the number of factors. *Econometrica* 81(3), 1203–1227.
- Alessi, L., M. Barigozzi, and M. Capasso (2010). Improved penalization for determining the number of factors in approximate factor models. *Statistics and Probability Letters* 80(23–24), 1806–1813.
- Anderson, T. W. (2003). *An Introduction to Multivariate Statistical Analysis* (3 ed.). New York: Wiley.
- Bai, J. and S. Ng (2002). Determining the number of factors in approximate factor models. *Econometrica* 70(1), 191–221.
- Bai, J. and S. Ng (2007). Determining the number of primitive shocks in factor models. *Journal of Business & Economic Statistics* 25(1), 52–60.
- Bai, J. and S. Ng (2008). Large dimensional factor analysis. *Foundations and Trends in Economics* 3(2), 89–163.
- Bates, B. J., M. Plagborg-Møller, 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.
- Breitung, J. and S. Eickmeier (2011). Testing for structural breaks in dynamic factor models. *Journal of Econometrics* 163(1), 71–84.
- Camacho, M., G. Perez-Quiros, and P. Poncela (2014). Extracting non-linear signals from several economic indicators. *Journal of Applied Econometrics*.
- Caner, M. and X. Han (2014). Selecting the correct number of factors in approximate factor models: The large panel case with group bridge estimators. *Journal of Business & Economic Statistics* 32(3), 359–374.
- Carrasco, M., L. Hu, and W. Ploberger (2014). Optimal test for markov switching parameters. *Econometrica* 82(2), 765–784.
- Chan, J. C. C., G. Koop, R. Leon-Gonzalez, and R. W. Strachan (2012). Time varying dimension models. *Journal of Business & Economic Statistics* 30(3), 358–367.
- Chauvet, M. (1998). An econometric characterization of business cycle dynamics with factor structure and regime switches. *International Economic Review* 39(4), 969–996.
- Chen, L., J. J. Dolado, and J. Gonzalo (2014). Detecting big structural breaks in large factor models. *Journal of Econometrics* 180(1), 30–48.

- Corradi, V. and N. R. Swanson (2014). Testing for structural stability of factor augmented forecasting models. *Journal of Econometrics* 182(1), 100–118.
- Diebold, F. and G. Rudebusch (1996). Measuring business cycle: A modern perspective. *Review of Economics and Statistics* 78(1), 67–77.
- Hamilton, J. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica* 57(2), 357–384.
- Hamilton, J. (1994). *Time Series Analysis*. Princeton, New Jersey: Princeton University Press.
- Hamilton, J. (2005). What’s real about the business cycle? *Federal Reserve Bank of St. Louis Review* 87(4), 435–452.
- Han, X. and A. Inoue (2015). Tests for parameter instability in dynamic factor models. *Econometric Theory FirstView*, 1–36.
- Hansen, B. (1992). The likelihood ratio test under non-standard conditions: Testing the Markov switching model of GNP. *Journal of Applied Econometrics* 7(S1), S61–S82.
- Harding, D. and A. Pagan (2002). Dissecting the cycle: A methodological investigation. *Journal of Monetary Economics* 49, 365–381.
- Kim, C.-J. (1994). Dynamic linear models with Markov-switching. *Journal of Econometrics* 60, 1–22.
- Kim, C.-J. and C. R. Nelson (1998). Business cycle turning points, a new coincident index, and tests of duration dependence based on a dynamic factor model with regime switching. *The Review of Economics and Statistics* 80(2), 188–201.
- Kim, C.-J. and C. R. Nelson (1999). Has the U.S. economy become more stable? A Bayesian approach based on a Markov-switching model of the business cycle. *Review of Economics and Statistics* 81, 608–616.
- McConnell, M. M. and G. Perez-Quiros (2000). Output fluctuations in the United States: What has changed since the early 1980’s? *The American Economic Review* 90(5), 1464–1476.
- McLachlan, G. and D. Peel (2000). *Finite Mixture Models* (1 ed.). New York: Wiley.
- Morley, J. and J. Piger (2012). The asymmetric business cycle. *The Review of Economics and Statistics* 94, 208–221.
- Onatski, A. (2010). Determining the number of factors from empirical distribution of eigenvalues. *The Review of Economics and Statistics* 92(4), 1004–1016.
- Sensier, S. and D. van Dijk (2004). Testing for volatility changes in U.S. macroeconomic time series. *The Review of Economics and Statistics* 86(3), 833–839.

- Stock, J. H. and M. W. Watson (1996). Evidence on structural instability in macroeconomic time series relations. *Journal of Business and Economic Statistics* 14(1), 11–30.
- Stock, J. H. and M. W. Watson (2002). Forecasting using principal components from a large number of predictors. *Journal of the American Statistical Association* 97, 1167–1179.
- Stock, J. H. and M. W. Watson (2005). Implications of dynamic factor models for VAR analysis. Working Paper 11467, NBER.
- Stock, J. H. and M. W. Watson (2009). Forecasting in dynamic factor models subject to structural instability. In J. Castle and N. Shephard (Eds.), *The Methodology and Practice of Econometrics: A Festschrift in Honour of David F. Hendry*, pp. 173–205. Oxford University Press.
- Stock, J. H. and M. W. Watson (2011). Dynamic factor models. In M. P. Clements and D. F. Hendry (Eds.), *Oxford Handbook of Forecasting*, Chapter 2, pp. 35–60. Oxford University Press.
- Yamamoto, Y. (2014). Forecasting with non-spurious factors in U.S. macroeconomic time series. Unpublished Manuscript.
- Yamamoto, Y. and S. Tanaka (2014). Testing for factor loading structural change under common breaks. Unpublished Manuscript.