Dating and Forecasting Turning Points by Bayesian Clustering with Dynamic Structure

Tom Rossa, Quentin Moayedpour, Naïl Khelifa

January 2025

Bayesian Statistics - ENSAE

Keywords: Time Series Forecasting - Bayesian Inference - Clustering - Monte Carlo Markov Chains
Project Github Repository

Contents

1 Introduction

Macroeconomics is a branch of economics that examines the structure, dynamics, and performance of an economy on a large scale. It relies on numerous aggregate indicators, such as GDP, unemployment rates, and inflation, as well as the policies that influence them. The article under study builds upon an extensive body of research dedicated to identifying shifts in macroeconomic cycles using these indicators. Since the foundational models developed in the 1940s and 1950s by Samuelson and Hicks, research has focused on understanding the mechanisms underlying economic fluctuations. During the 1970s and 1980s, advances such as the unobservable component models proposed by Harvey and Clark (1977) and the Markov Switching models introduced by Hamilton (1989) significantly enhanced the modeling of transitions between periods of growth and recession. By 2008, these efforts had been further enriched by the adoption of Bayesian methods and clustering techniques, enabling more precise detection of economic cycles from a wide array of macroeconomic indicators.

In the article by [4], a methodology for detecting business cycle turning points is introduced, leveraging the information embedded in time series data that represent macroeconomic indicators. This idea draws inspiration from the work of [3], which proposed a clustering approach for time series under the assumption of a statistical model. The methodology detailed here operates within the framework of Bayesian inference, modeling the parameters of the time series, their assignment to different clusters and the dynamic structure of interactions between clusters. The primary objective is to estimate these parameters as accurately as possible to predict turning points in cycles and to enhance the forecasting of the time series themselves.

2 Theoretical Framework

In this section, we introduce the statistical model used to represent our dataset of economic time series. This model is broad and flexible enough to account for regime switches within economic cycles, clustering of time series, inter-cluster dynamics, as well as heterogeneity within clusters. It is worth noting that, as a preliminary step, the time series are centered, scaled, and detrended to mitigate potential scale effects and biases during the inference process.

2.1 Model

In general, we consider a dataset consisting of N time series observed over $t \in \{1, \cdots, T\}$. This dataset is denoted as $\mathcal{D}_T = \{Y_T, Y_{T-1}, \cdots, Y_1\}$, where $Y_t = \{y_{1t}, y_{2t}, \cdots, y_{Nt}\}$. Each time series, representing the growth rates of various macroeconomic indicators, is modeled using an autoregressive process of order p ($AR(p)$):

$$y_{it} = \mu^i_{I_{it}} + \phi^i_1 y_{i,t-1} + \cdots + \phi^i_p y_{i,t-p} + \varepsilon_{it},$$

where $\varepsilon_{it} \overset{\text{iid}}{\sim} N(0, \sigma^2/\lambda_i)$ is white noise with variance specific to the series but dependent on the cluster to which it belongs. To capture the transitions between periods of economic expansion and recession, an unobservable state indicator I_{it} is introduced, taking values in $\{1, 2\}$. We define state 1 as periods of below-average growth and state 2 as periods of above-average growth, such that:

$$\mu_i = \begin{cases} \mu_{i1} & \text{if } I_{it} = 1, \\ \mu_{i2} & \text{if } I_{it} = 2, \end{cases} \quad \text{with } \mu_{i1} \leq 0 \text{ and } \mu_{i2} > 0.$$

Following the works of [1] and extended in [5], the state indicator I_{it} is assumed to follow a Markov switching process:

$$\mathbb{P}(I_{it} = l \mid I_{it-1}, \cdots, I_{i1}) = \mathbb{P}(I_{it} = l \mid I_{it-1} = j) = \xi^k_{jl}, \quad \forall i, j \in \{1, 2\}.$$

The time series are assumed to be stationary, meaning that the autoregressive polynomial ϕ has no unit roots.

2.2 Clusters & Dynamic Structure

Beyond identifying turning points in business cycles, the model also aims to cluster the time series. To this end, we introduce the indicator S_i, which assigns each series i to one of K groups. This indicator is modeled using a multinomial logit to incorporate prior information into the estimation of group probabilities. Specifically:

$$P(S_i = k \mid \gamma_1, \dots, \gamma_{K-1}, \gamma_{z1}, \dots, \gamma_{z,K-1}) = \frac{\exp(\gamma_k + Z_i \gamma_{zk})}{1 + \sum_{l=1}^{K-1} \exp(\gamma_l + Z_i \gamma_{zl})}, \quad k = 1, \dots, K-1,$$

where Z_i is a vector of series-specific features that influence group classification, with γ representing unknown, group-specific parameters. In practice, it is recommended to determine the prior group probabilities based on the correlation of each series with GDP and industrial order books. The model assumes a dynamic relationship where the second group ($S_i = 2$) leads the first group ($S_i = 1$). This implies that when the series in group 2 switch phases, the series in group 1 are certain to switch within the same cycle. A global state variable I_t^* captures the $J^* = 4$ possible combinations of the group states:

$$I_t^* = \begin{cases} 1, & (I_{1t} = 1, I_{2t} = 1), \\ 2, & (I_{1t} = 1, I_{2t} = 2), \\ 3, & (I_{1t} = 2, I_{2t} = 1), \\ 4, & (I_{1t} = 2, I_{2t} = 2). \end{cases}$$

The transition matrix ξ^* enforces the leading role of group 2 by restricting certain transitions to zero:

$$\xi^* = \begin{bmatrix} \xi_{1,1}^* & \xi_{1,2}^* & 0 & 0 \\ 0 & \xi_{2,2}^* & 0 & \xi_{2,4}^* \\ \xi_{3,1}^* & 0 & \xi_{3,3}^* & 0 \\ 0 & 0 & \xi_{4,3}^* & \xi_{4,4}^* \end{bmatrix}.$$

More generally, the assignment of groups 1 and 2 is determined by the variable ρ^*, which takes its values in the set of permutations of $\{1, 2, 0_{K-2}\}$. This variable designates group 2 as the leading group, group 1 as the coincident group, and all other groups as independent of the leading group's characteristics. If the dynamic structure is unknown, ρ^* assigns an equal prior probability to each of the $L = K(K-1)$ possible permutations.

The purpose of this model is to group time series with similar dynamics and to capture the interactions between these groups. A classical approach in panel data modeling involves pooling across groups, where parameters are estimated as identical for all groups or individuals without distinction. However, this method strikes a balance between complete pooling and total independence of parameters. Specifically, parameters are estimated within a hierarchical framework: each group is assigned a specific mean, and individual group parameters are adjusted toward this mean while preserving heterogeneity within groups. This approach helps mitigate the risk of overfitting and enhances the statistical robustness of the inference compared to independently estimating parameters for each series. Additionally, it enables the model to capture intergroup variations while leveraging shared information through the dynamic structure.

- **Group-Specific Mean Parameters**: For a time series y_i belonging to group k and state $I_{k,t} = l \in \{1, 2\}$, the mean parameter μ_i is drawn from a normal distribution:

  $$\mu_i \sim \mathcal{N}(\mu_l^k, q_k^2),$$

  where q_k^2 is the variance, capturing the heterogeneity within the group.

- **Autoregressive Parameters:** The autoregressive coefficients for series i within group k are drawn from a group-specific multivariate normal distribution:

$$(\phi_i^{(1)}, \ldots, \phi_i^{(p)}) \sim \mathcal{N}(\phi_k, Q_k),$$

  where: ϕ_k is the group-specific mean vector of autoregressive parameters and Q_k is the covariance matrix, controlling the variability and correlation structure of the autoregressive parameters within the group.

After introducing all the parameters of the model, the likelihood of our dataset of time series can be expressed as:

$$\begin{aligned}
\mathcal{L}(Y^T|\psi) &= \prod_{t=p+1}^{T} \prod_{i=1}^{N} f(y_{it}|y_{i,t-1:1}, \mu^{S_i}_{I_{S_i t}}, \phi^{S_i}, Q^{S_i}, \lambda_i, \sigma^2) \\
&= \prod_{t=p+1}^{T} \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi\nu_{it}^{S_i}}} \exp\left(-\frac{1}{2\nu_{it}^{S_i}}\left(y_{it} - \mu^{S_i}_{I_{S_i t}} - \sum_{j=1}^{p} \phi_j^{S_i} y_{i,t-j}\right)^2\right)
\end{aligned}$$

where $\nu_{it}^{S_i}$ is the variance parameter of the conditional distribution of the observations. This variance encompasses both the variability induced by the residuals of the autoregressive process, $(\varepsilon_{it})_{it}$, and the variance arising from the heterogeneity of time series within the same group, specifically due to the autoregressive parameters.

3 Bayesian Inference

All the parameters of the previously introduced statistical model are encapsulated in a parameter vector θ, which is extended to $\psi = (\theta, S^N, I^T, \lambda^N, \rho^*)$. This augmented parameter set incorporates information on the different clusters, the phases of the cycle, and the dynamic structure between clusters. To perform statistical inference, we adopt a Bayesian framework. The prior distribution, denoted as $\pi(\psi)$, encodes our prior beliefs about the possible values of the parameters, which will be updated using the observed data.

3.1 Choice of the Prior Distribution

When considering a known dynamic structure, ρ^*, the prior distribution takes the following form:

$$\pi(\psi) = \pi(I_T^* \mid \rho^*, \xi^*) \prod_{\rho^*(k)=0} \pi(I_k^T \mid \rho^*, \xi^k)\pi(S^N \mid \gamma, \gamma_z, Z^N)\pi(\lambda^N)\pi(\theta)$$

The prior distribution is therefore defined in a block-wise manner, with the density $\pi(\theta)$ further subdivided into multiple parameter blocks. In cases where the dynamic structure is not known, an additional step can be incorporated into the modeling and inference process to estimate it from the data. This framework allows the incorporation of prior beliefs into the distribution in various ways. First, the dynamic structure can be fixed, for example, as $\rho^* = (1, 2, 0_{K-2})$. Based on this structure, it is possible to pre-assign certain time series to specific groups according to prior beliefs, as illustrated in the referenced study. For instance, GDP and its main components (such as consumption and investment) are classified into the coincident group. Leading indicators (such as expectations about orders, recent production, order books, and industrial sector confidence surveys) are classified into the leading group. Furthermore, prior probabilities for a time series belonging to a particular group are adjusted based on its correlation with two reference variables:

- GDP, for the coincident group.

- Industrial orders, for the leading group.

This approach ensures the integration of meaningful prior information, which enhances the model's interpretability and its alignment with domain-specific knowledge.

3.2 Sampling the Posterior Distribution via MCMC

To conduct Bayesian inference, it is necessary to update the various parameters by sampling from the posterior distribution, given by

$$\pi(\psi|Y^T) \propto \mathcal{L}(Y^T|\psi) \cdot \pi(\psi)$$

The sampling procedure which aims to draw from the posterior distribution $\pi(\psi|Y^T)$ uses a structured approach that accounts for the hierarchical dependencies in the model.

- **Sampling Group Indicators** $\pi(S^N|Y^T, I^T, \rho^*, \lambda^N, \theta)$ **:** The group indicator for each time series is sampled individually, and depends on the observations and the prior information contained in the variable Z_i.
- **Sampling State Indicators** $\pi(I^T|Y^T, S^N, \rho^*, \lambda^N, \theta)$**:** The group-specific state indicators are drawn using multi-move sampling, which leverages the encompassing specification I^* to jointly sample the state indicators for the coincident and leading groups.
- **Sampling Series-Specific Weights** $\pi(\lambda^N|Y^T, S^N, I^T, \lambda^N, \theta)$**:** The weights associated with each time series are sampled independently from Gamma distributions.
- **Sampling Model Parameters** $\pi(\theta|Y^T, S^N, I^T)$**:** Most posterior distributions are conjugate to their priors, enabling efficient sampling. However, for certain parameters (e.g., γ and γ_z), which influence group probabilities, the posterior is not in closed form. These parameters are sampled using a Metropolis-Hastings step.

4 Experiments

We apply the model to Austrian quarterly data from 1988 to 2003. The dataset includes several economic time series, such as GDP, investment indices, consumption indices, and industrial production. The objective is to estimate and classify these series into $K = 3$ distinct groups: series that lead the business cycle ($S_i = 2$), series that lag behind it ($S_i = 1$), and independent series ($S_i = 3$). Prior to analysis, the time series are detrended and standardized. An example showcasing three series from the dataset is presented in Figure 1.

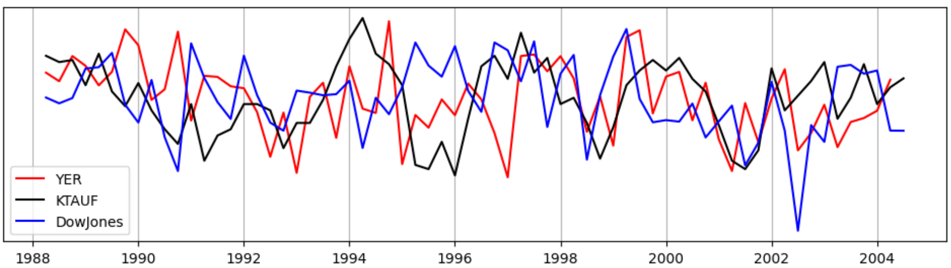
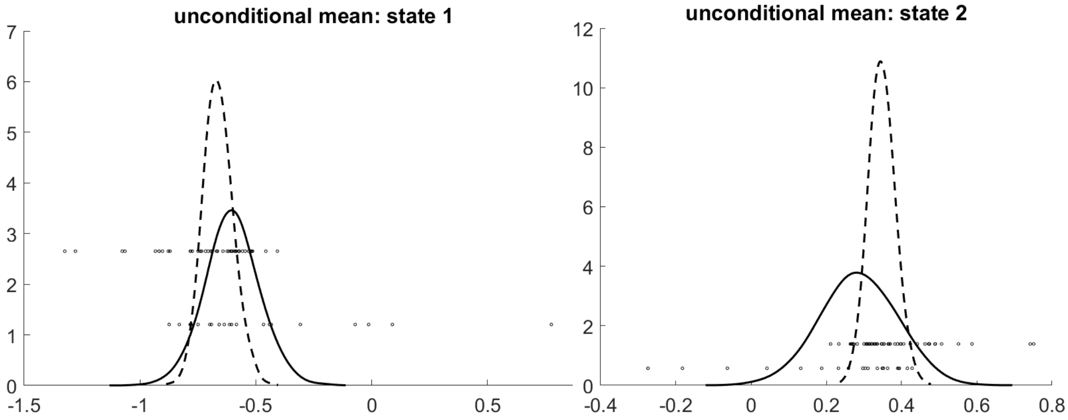


Figure 1: Detrended and scaled series

Additionally, we impose a constraint on the model such that the variable GDP belongs to group 1 ($S_{i,GDP} = 1$), and is therefore no longer a parameter to be estimated, but a fixed parameter within the model. Figure 4 illustrates the significance of accounting for the heterogeneity of the series based on the assigned group. It shows that the sum of the AR coefficients varies significantly across groups. The characterization by AR coefficients is crucial because, as shown in Figure 2, the series do not appear to be easily distinguishable based on their (unconditional) mean. This highlights the importance of the assumed statistical model in classifying the series.



(a) State 1 (b) State 2

Figure 2: Unconditionnal mean in different states

The graph 2 represents the unconditional average of each time series group in both states, depending on whether they are below or above the average growth. This shows that our two series groups have rather similar behaviors since the series are only slightly offset from each other. It is worth noting that they are more offset in state 2.

The graph 3 illustrates the effect of the correlation with GDP and the order book between groups 1 (figure 4a) and 2 (figure 4b). The curve represents:

$$\mathbb{P}(S = j \mid Z, Y^T) = \int \int \mathbb{P}(S = j \mid Z, \gamma, \gamma_z) \pi(\gamma, \gamma_z \mid Y^T) d\gamma d\gamma_z,$$

the conditional probability on GDP and the order book (KTAUF variable). This probability is highest for the series in group 1, which are strongly correlated with GDP. The correlation with GDP seems to play a more important role since the probability varies more according to the value of GDP. Here we find the definition of the first group of variables, which are coincident with the cycle and hence correlated with GDP. For group 2, the probability value is maximal for series that are strongly correlated with the order book, and this variable appears to be more important (with GDP correlation having little influence).

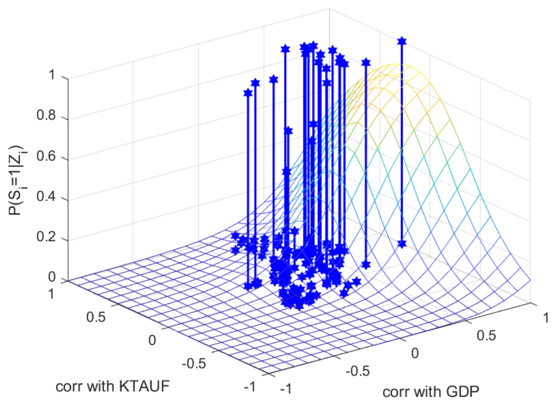
The height of the vertical lines in graph 3 represents the a posteriori probability

$$\mathbb{P}(S_i = j \mid Z_i, Y^T)$$

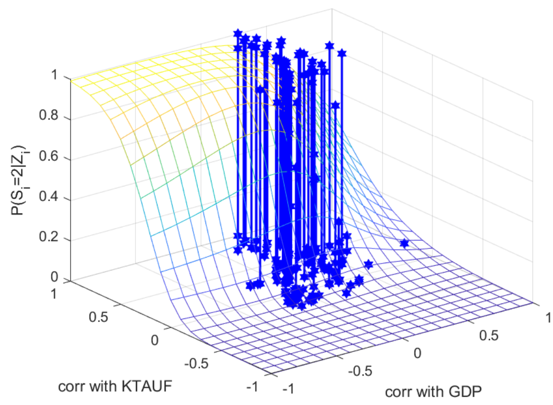
for a series to belong to groups 1 or 2. In the vast majority of cases, this probability was updated to 0 or 1, thanks to the additional information obtained. This allows us to verify the quality of our pre-classification: it seems consistent because the series classified in the coincident group have a high a posteriori probability of belonging to that group, and similarly for the series classified in the advanced group.

The proposed model provides a robust approach for time series classification, incorporating prior knowledge about their dynamics and their relationships with the economic cycle. Unlike traditional machine learning models, which are often generic, this Bayesian approach guides the estimation of parameters and reduces uncertainty, particularly for key series such as GDP (group 1).

To evaluate this approach, we compared the results with those from a standard machine learning model, ASTRIDE ([2]). The classifications obtained differ significantly: the Bayesian model captures nuances specific to the series, clearly distinguishing those that align with the



(a) Group 1



(b) Group 2

Figure 3: Prior and posterior probabilities $\mathbb{P}(S = j \mid Z, Y^T)$ and $\mathbb{P}(S_i = j \mid Z_i, Y^T)$.

economic cycle (such as GDP) from those that lead it (such as order books), whereas the machine learning model produces more generic results (Table 1 & 2). The classification results remain notably different between the two tables.

These differences highlight the strength of the Bayesian approach: it integrates expert knowledge and structural assumptions, producing unique and interpretable results. Its ability to model transitions between economic phases and handle the heterogeneity of the series makes it a valuable tool for economic analysis, surpassing machine learning models in this specific context.

5 Conclusion

This study has examined a Bayesian framework for classifying time series and identifying turning points in economic cycles, as proposed in the original paper. The approach leverages prior knowledge and a dynamic clustering structure to capture the nuanced relationships between macroeconomic indicators, such as GDP and industrial order books. By integrating expert insights and structural assumptions, the Bayesian model provides robust and interpretable results, making it a powerful tool for understanding economic cycles. The model, however, relies on strong statistical assumptions about the time series, which may limit its applicability in certain contexts.

A critical limitation is that the model requires a deep understanding of the time series to select appropriate priors, ensuring reliable estimations. Additionally, the model assumes that all time series must have the same length and observations at identical time points. This constraint typically necessitates the use of quarterly data, which often results in datasets with a limited number of observations. As a result, an inappropriate choice of prior can prevent the model from converging to accurate estimates, as the limited data availability increases the sensitivity to prior specifications.

References

[1] P. Bengoechea and G. Pérez-Quiros. *A useful tool to identify recessions in the euro area.* Economic Papers 215. European Economy, European Commission, Directorate-General for Economic and Financial Affairs, 2004.

[2] Sylvain W. Combettes, Charles Truong, and Laurent Oudre. *ASTRIDE: Adaptive Symbolization for Time Series Databases.* 2023. arXiv: `2302.04097 [cs.LG]`. URL: `https://arxiv.org/abs/2302.04097`.

[3] Sylvia Frühwirth-Schnatter and Sylvia Kaufmann. "Model-Based Clustering of Multiple Time Series". In: *Journal of Business & Economic Statistics* 26.1 (2008), pp. 78–89.

[4] Sylvia Kaufmann. "Dating and forecasting turning points by Bayesian clustering with dynamic structure: a suggestion with an application to Austrian data". In: *Journal of Applied Econometrics* 25.2 (2010), pp. 309–344.

[5] R. Paap, R. Segers, and D. van Dijk. *Do leading indicators lead peaks more than troughs?* Econometric Institute Report EI 2007-08. Econometric Institute and Tinbergen Institute, Erasmus University Rotterdam, 2007.

Annexe

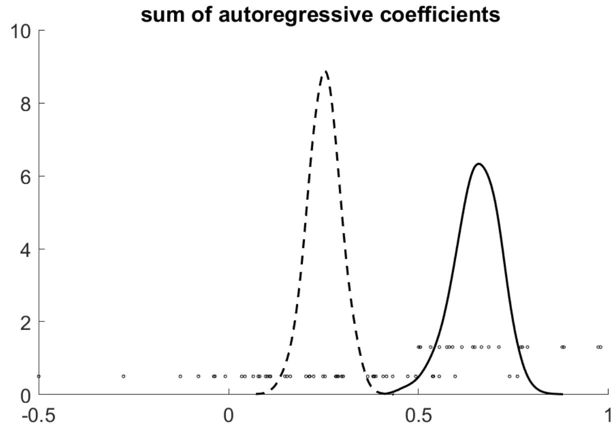
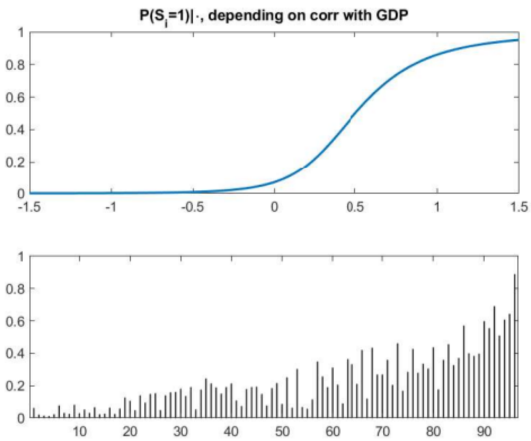
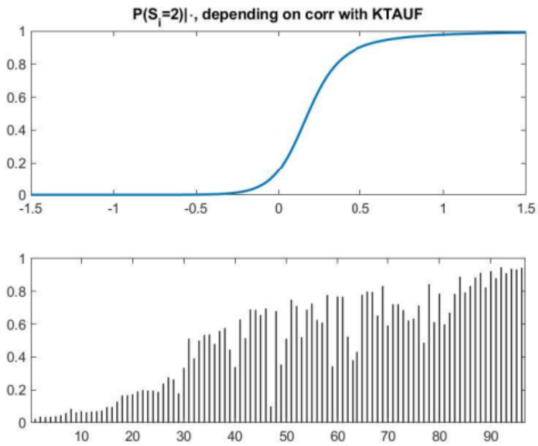


Figure 4: Sum of autoregressive coefficients
Straight lines: groupe 1, Pointed: groupe 2



(a) GDP & coincident group



(b) KTAUF & leading group

Figure 5: Impact of the correlation with GDP/Order Books on the probability to belong to the coincident/leading group

The bar represent the posterior probability to belong to the group j: $P(S_i = j|Z_i, Y^t)$ for $j = 1, 2$ and $t = 1, \cdots T$.

Table 1: Series classification Bayesian model

$S_i = 1$	$S_i = 2$	$S_i = 3$
YER	EEN	TOT
PCR	QTAUF	USD
ITR	QTEXPA	HICP
GCR	QTLAG	HICP$_FO$
MTR	QTPR	HICP$_UF$
XTR	QTPRO	HICP$_G$
TLIARG86	QTBAUF	HICP$_IG$
TLIANG86	QTBPR	HICP$_GX$
GHPIOS	QTBBGL	HICP$_XA$
EXPG	QTBAGL	HICP$_XE$
EXP7	KTPROL	HICP$_XG$
EXP8	KTAUF	VPIG86
IMPG	KTAUSL	VPI$_WOH8$
IMP6	KTLAG	TLIG86
IMP7	KTPRON	GHPIGG
EXP$_US$	KTVPN	GHPIGL
EXP$_EU$	BAUVPN	GHPIKONG
EXP$_DE$	EECOS	GHPIINVG
IMP$_US$	EINDSE	IMP8
IMP$_EU$	EBAUSE	M2
IMP$_DE$	EHANSE	M3
	EKONSE	DowJones
	IFOERW	YIELD
	IFOKL	DCR$_HH$
	IFOGL	DCR$_F$
	PMI	DEBT
	HICP$_PF$	DCR
	HICP$_E$	
	HICP$_S$	
	HICP$_XF$	
	GHPIG	
	GHPIGK	
	GHPIVBG	
	GHPIINTG	
	OEL	
	EXP6	
	ALQNSA	
	ALOSM	
	ALOSW	
	OFST	
	STANDR	
	INDPROD	
	ATX	
	M1	
	DAX	
	STI	
	SEKMRE	
	DCR$_G$	

Table 2: Series classification ASTRIDE

$S_i = 0$	$S_i = 1$	$S_i = 2$
QTBAUF	YER	QTAUF
TLIG86	PCR	QTEXPA
TLIARG86	ITR	QTPR
TLIANG86	GCR	QTPRO
GHPIINTG	MTR	QTBPR
OFST	XTR	QTBBGL
STANDR	TOT	QTBAGL
M1	EEN	KTPROL
DowJones	USD	KTAUF
DEBT	QTLAG	KTAUSL
DCR	HICP	KTLAG
	VPIG86	KTPRON
	GHPIG	KTVPN
	GHPIOS	BAUVPN
	GHPIGG	EECOS
	GHPIGL	EINDSE
	GHPIGK	EKONSE
	GHPIVBG	IFOERW
	GHPIKONG	IFOKL
	GHPIINVG	IFOGL
	OEL	PMI
	EXP7	ALOSM
	EXP8	ALOSW
	IMPG	INDPROD
	IMP6	ATX
	IMP7	SEKMRE
	IMP8	
	ALQNSA	
	EBAUSE	
	EHANSE	
	M2	
	M3	
	DAX	
	STI	
	YIELD	