

## Outliers in dynamic factor models

Roberto Baragona, *Dept. Sociology & Communication, Sapienza Univ. of Rome, Italy*  
Francesco Battaglia, *Dept. Statistics, Sapienza Univ. of Rome, Italy*

### Abstract

Dynamic factor models have a wide range of applications in econometrics and applied economics. The basic motivation resides in their capability of reducing a large set of time series to only few indicators (factors). If the number of time series is large compared to the available number of observations then most information may be conveyed to the factors. This way low dimension models may be estimated for explaining and forecasting one or more time series of interest. It is desirable that outlier free time series be available for estimation. In practice, outlying observations are likely to arise at unknown dates due, for instance, to external unusual events or gross data entry errors. Several methods for outlier detection in time series are available. Most methods, however, apply to univariate time series while even methods designed for handling the multivariate framework do not include dynamic factor models explicitly. A method for discovering outliers occurrences in a dynamic factor model is introduced that is based on linear transforms of the observed data. Some strategies to separate outliers that add to the model and outliers within the common component are discussed. Applications to simulated and real data sets are presented to check the effectiveness of the proposed method.

AMS 2000 subject classifications: Primary 62M10, 62H25; secondary 62P20.

Keywords: Dynamic factor models, Multivariate time series, Outliers.



Full Text: [PDF](#)

Baragona, Roberto, Battaglia, Francesco, Outliers in dynamic factor models, *Electronic Journal of Statistics*, 1, (2007), 392-432 (electronic). DOI: 10.1214/07-EJS082.

## References

- [1] Anderson, T. W. (1958). *An Introduction to Multivariate Statistical Analysis*. John Wiley & Sons, New York. [MR0091588](#)
- [2] Box, G. E. P. and Tiao, G. C. (1977). A canonical analysis of multiple time series. *Biometrika* 64, 355–365. [MR0519089](#)
- [3] Brillinger, D. R. (1981). *Time Series, Data Analysis and Theory* (expanded ed.). Holden-Day, San Francisco. [MR0595684](#)
- [4] Chang, I., Tiao, G. C., and Chen, C. (1988). Estimation of time series parameters in the presence of outliers. *Technometrics* 30, 193–204. [MR0943602](#)
- [5] Forni, M., Hallin, M., Lippi, M. and Reichlin, L. (2000). The generalized dynamic-factor model: identification and estimation. *The Review of Economics and Statistics* 82, 540–554.
- [6] Forni, M. and Reichlin, L. (1998). Let's get real: a dynamic factor analytical approach to disaggregated business cycle. *Review of Economic Studies* 65, 453–474.
- [7] Fox, A. J. (1972). Outliers in time series. *Journal of the Royal Statistical Society*,

- [8] Galeano, P., Peña, D., and Tsay, R. S. (2006). Outlier detection in multivariate time series by projection pursuit. *Journal of the American Statistical Association* 101, 654–669. [MR2256180](#)
- [9] Geweke, J. (1977). The dynamic factor analysis of economic time-series models. In *Latent Variables in Socio-Economic Models*, D. J. Aigner and A. S. Goldberger, Eds. North-Holland, New York, 365–383.
- [10] Giri, N. (1965). On the complex analogues of  $T^2$  and  $R^2$  tests. *The Annals of Mathematical Statistics* 36, 664–670. [MR0182088](#)
- [11] Johnson, R. A. and Wichern, D. W. (1998). *Applied Multivariate Statistical Analysis*. Prentice Hall, Upper Saddle River, New Jersey.
- [12] Lütkepohl, H. (1996). *Handbook of Matrices*. John Wiley & Sons, Chichester. [MR1433592](#)
- [13] Magnus, J. R. and Neudecker, H. (1988). *Matrix Differential Calculus with Applications in Statistics and Econometrics*. John Wiley & Sons, Chichester. [MR940471](#)
- [14] Peña, D. and Box, G. E. P. (1987). Identifying a simplifying structure in time series. *Journal of the American Statistical Association* 82, 836–843. [MR909990](#)
- [15] Peña, D. and Prieto, F. J. (2001). Multivariate outlier detection and robust covariance matrix estimation. *Technometrics* 43, 286–310. [MR1943185](#)
- [16] Press, J. S. (1972). *Applied Multivariate Analysis*. Holt, Rinehart and Winston. [MR0420970](#)
- [17] Priestley, M. B. (1981). *Spectral Analysis and Time Series*. Academic Press, New York. [MR628736](#)
- [18] Quenouille, M. H. (1957). *The Analysis of Multiple Time Series*. Charles W. Griffin, London. [MR0093092](#)
- [19] Sargent, T. J. and Sims, C. A. (1977). Business cycle modeling without pretending to have too much a priori economic theory. In *New Methods in Business Cycle Research: Proceedings from a Conference*, C. A. Sims, Ed. Federal Reserve Bank of Minneapolis, Minneapolis, MN, 45–110.
- [20] Stock, J. H. and Watson, M. W. (1998). Diffusion indexes. Technical Report 6702, National Bureau of Economic Research. <http://www.nber.org/papers/w6702>.
- [21] Stock, J. H. and Watson, M. W. (2002). Forecasting using principal components from a large number of predictors. *Journal of the American Statistical Association* 97, 1167–1179. [MR1951271](#)
- [22] Stock, J. H. and Watson, M. W. (2004). Combination forecasts of output growth in a seven-country data set. *Journal of Forecasting* 23, 405–430.
- [23] Tsay, R. S., Peña, D., and Pankratz, A. E. (2000). Outliers in multivariate time series. *Biometrika* 87, 789–804. [MR1813975](#)
- [24] Yeredor, A. (2002). Non-orthogonal joint diagonalization in the least-squares sense with application in blind source separation. *IEEE Transactions on Signal Processing* 50, 1545–1553. [MR1931239](#)
- [25] Ziehe, A. and Müller, K.-R. (1998). Tdsep — an efficient algorithm for blind separation using time structure. In *Proceedings of the International Conference on ICANN, Perspectives in Neural Computing*. Springer, Berlin, 675–680.

