Time-varying Factor-augmented Regression Model for High-dimensional Data

Wei Liu, Ph.D. Candidate
Department of Economics, Louisiana State University

Overview

This paper introduces smooth time-varying coefficients to the standard factor-augmented regression (FR) model for high-dimensional data.

Using time-varying coefficients in regression effectively addresses parameter instability and is commonly applied in macroeconomic and financial studies.

This paper shows that consistency of estimated forecasting coefficients to the true forecasting coefficients and the feasible forecast $\hat{Y}_{T+h|T}$ asymptotically converges to the optimal infeasible forecast $Y_{T+h|T}$.

The novel forecasting method captures the structural change of parameters in simulations and empirical applications.

Research Idea

1. Both factor loadings and regression coefficients are time-varying by using a local PCA procedure
2. Employs a boundary kernel (Hong and Li, 2005; Su and Wang, 2017) approach to address boundary bias
3. Wei and Zhang (2020) have explored a comparable method, but they failed to provide theoretical justification
4. Contribution: bridge the theoretical gap concerning the time-varying FR model

Standard FR Model

$$x_{it} = \phi_i' F_t + e_{it}$$

$$y_{t+h} = \alpha' F_t + \beta' W_t + \eta_{t+h}$$

Time-varying Version FR Model

$$x_{it} = \phi_{it}' F_t + e_{it}$$

$$y_{t+h} = \alpha_t{}' F_t + \beta_t' W_t + \eta_{t+h}$$

After transformation with local nonparametric method

$$y_{t+h} = \alpha_t'(\tau_t)[H^{(T)\prime}]^{-1} \times K_{b,tT}^{-1/2}\left[H^{(T)\prime}\right] F_t^{(T)} \beta_t' W_t + \eta_{t+h}$$

Figure 1. Standard FR Model Vs. Time-varying FR Model

Asymptotic Properties

The last equation on the right side of Figure 1 can be expressed in the following compact form:

$$y_{T+h|T} = \rho^{(T)\prime}(1) G_t^{(T)}$$

Theorem 1

Suppose that all assumptions and conditions hold, if $\frac{\sqrt{T}B^{-1}}{N} \to 0$, then

$$\sqrt{T}B^{-1}\left[\hat{\rho}^{(T)\prime}(1) - \rho^{(T)\prime}(1)\right] \to \mathrm{N}(0, \Sigma_\delta)$$

where Σ_δ is a matrix of variance, also define $B \equiv \begin{matrix} \frac{1}{\sqrt{b}} & 0 \\ 0 & I \end{matrix}$, b is the bandwidth using in the local nonparametric method.

Theorem 2

Based on the assumptions of Theorem 1 and $\sqrt{N}/T \to N(0,1)$,

$$\frac{\hat{y}_{T+h|T} - y_{T+h|T}}{B\sqrt{Var(\hat{y}_{T+h|T})}} \xrightarrow{d} N(0,1)$$

where $Var(\hat{y}_{T+h|T})$ is defined following Bai and Ng's (2006) paper.

Empirical Applications

Target Variables: 8 macroeconomic aggregates (real personal income less transfers, total industrial production, real manufacturing and trade sales, the number of employees on nonagricultural payrolls, the consumer price index (CPI), the CPI less food, the producer price index (PPI), and the real personal consumption expenditure)

Data: Well-known Stock and Watson (2002)
Quarterly data is from 1990 to 2019; the out-of-sample forecast period is from 1987 to 2019

The exogenous predictors selected in this paper are 33 in total; They mainly involve asset prices, measures of real economic activity, price indices, and monetary measures.

Benchmarks to compare the prediction ability: the standard FR model, time-varying FR model but using rolling window with a size of 40 or 80.

The criterion for evaluating prediction ability is the root mean squared forecasting error (RMSFE, as shown in Table 3). A smaller RMSFE indicates better prediction ability. Table 4 also provides the Giacomini and White's equal conditional predictive ability tests. The positive test statistics means the proposed model has better conditional predictive ability than the benchmarks.

Empirical Results

Table 3: RMSFEs for Eight Macroeconomic Variables

	RMSFE (h=1)				RMSFE (h=2)			
	TVFR	FR	RW40	RW80	TVFR	FR	RW40	RW80
Real personal income less transfer receipts	0.0081*	0.0085	0.0090	0.0088	0.0056*	0.0062	0.0059	0.0061
Total industrial production	0.0094*	0.0116	0.0102	0.0098	0.0075*	0.0081	0.0080	0.0076
Real manufacturing & trade industries sales	0.0111	0.0155	0.0109*	0.0109*	0.0073	0.0104	0.0071	0.0072
Total number of nonagricultural employees	0.0074*	0.0075	0.0084	0.0081	0.0053*	0.0056	0.0057	0.0055
CPI	0.0028*	0.0032	0.0029	0.0028	0.0017*	0.0036	0.0018	0.0017*
CPI less food	0.0032*	0.0039	0.0033	0.0032*	0.0019*	0.0042	0.0020	0.0020
PPI for finished goods	0.0068*	0.0132	0.0072	0.0070	0.0038*	0.0046	0.0039	0.0039
Real personal consumption expenditure	0.0090*	0.0133	0.0096	0.0097	0.0063*	0.0086	0.0067	0.0066

Notes: TVFR: forecasts use the time-varying factor-augmented regression model; FR: forecast use the standard factor-augmented regression model; RW40: forecasts use a rolling window with a size of 40; RW80: forecasts use a rolling window with a size of 80. "*" denotes the lowest RMSFE value. The RMSFEs for forecasts using a rolling window of size 60 are omitted, as most values are same those of RW40 or RW80. This omission does not affect the demonstrated superiority of the TVFR model's forecasting ability. The RMSFEs for forecasts using a rolling window of size 60 are omitted, as most values are the same as those of RW40 or RW80. This omission does not affect the demonstrated superiority of the TVFR model's forecasting ability.

Table 4: The Conditional Predictive Ability Test for Eight Macroeconomic Variables

	RMSFE (h=1)			RMSFE (h=2)		
	FR	RW40	RW80	FR	RW40	RW80
Real personal income less transfer receipts	17.59	10.40	14.34	55.65	20.50	21.71
	(0.0000)	(0.0055)	(0.0007)	(0.0000)	(0.0000)	(0.0000)
Total industrial production	5.55	0.14	0.31	18.62	6.32	9.25
	(0.0623)	(0.9313)	(0.8556)	(0.0000)	(0.0424)	(0.0098)
Real manufacturing & trade industries sales	0.01	1.87	2.36	9.18	4.82	9.08
	(0.0005)	(0.3928)	(0.3070)	(0.0101)	(0.0898)	(0.0107)
Total number of nonagricultural employees	57.70	28.63	58.61	60.05	30.89	55.11
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
CPI	-2.74	3.50	-9.20	-1.48	-0.69	-1.29
	(-0.2535)	(0.1736)	(-0.0100)	(-0.4783)	(-0.7084)	(-0.5237)
CPI less food	4.65	3.24	5.99	-1.47	-0.78	-1.49
	(0.0976)	(0.1981)	(0.0501)	(-0.4792)	(-0.6758)	(-0.4748)
PPI for finished goods	2.48	1.21	2.74	4.73	4.15	5.31
	(0.2896)	(0.5460)	(0.2541)	(0.0939)	(0.1245)	(0.0703)
Real personal consumption expenditure	0.05	2.40	3.29	36.02	2.96	6.80
	(0.8354)	(0.3015)	(0.1926)	(0.0000)	(0.2272)	(0.0334)

Notes: The entries are the equal conditional predictive ability test statistics when comparing the benchmark models to the TVFR methods. The numbers within parentheses are p-values of the tests.

Conclusions and Discussions

Empirical results demonstrate that the proposed model outperforms the benchmark methods in terms of forecast accuracy.

This paper will further explore the estimation of forecasting intervals in the future. I believe that conducting only point forecasting is insufficient; constructing confident forecasting intervals could be more useful and important.

References

1. Bai, J., & Ng, S. (2006). Confidence intervals for diffusion index forecasts and inference for factor-augmented regressions. *Econometrica, 74(4), 1133-1150.*
2. Hong, Y., & Li, H. (2005). Nonparametric specification testing for continuous-time models with applications to term structure of interest rates. *The Review of Financial Studies, 18(1), 37-84.*
3. Su, L., & Wang, X. (2017). On time-varying factor models: Estimation and testing. *Journal of Econometrics, 198(1), 84-101.*
4. Stock, J. H., & Watson, M. W. (2002). Forecasting using principal components from a large number of predictors. *Journal of the American statistical association, 97(460), 1167-1179.*
5. Wei, J. and Zhang, Y. (2020). A time-varying diffusion index forecasting model. *Economics Letters, 193:109337*.

Contact

Wei Liu
Department of Economics
Louisiana State University
Email: wliu27@lsu.edu
Website: https://sites.google.com/view/wei-liu-econ/