Nowcasting GDP with targeted predictors: A model averaging for dynamic factor models

Yongmiao Hong[1,2,3] , Yizhou Kuang[4] , Yuying Sun[1,2,3]

[1]Academy of Mathematics and Systems Science, Chinese Academy of Sciences
[2]Center for Forecasting Science, Chinese Academy of Sciences
[3]School of Economics and Management, University of Chinese Academy of Sciences
[4]Department of Economics, Cornell University

July 11, 2021

Outline

1. Motivation
2. Unbalanced Data Structure
3. Methodology
4. Asymptotic Theory
5. Simulation
6. Application
7. Conclusion

Motivation

- Real-time Gross domestic production (GDP) prediction, along with some other important policy indicators, has drawn an increasing attention during the last decades.
- Many policy decisions, including monetary policy, need to be made in real time and are based on assessments of current and future economic conditions.
- In central banks, estimated current-quarter GDP figures are often used as relevant inputs for model-based longer-term forecasting exercises.

Motivation (cont'd)

Some challenges:

- How to bridge information contained in monthly data with the quarterly GDP (Mixed frequency issue)?
  - ✓ Bridge equations (Baffigi et al., 2004; Kitchen and Monaco, 2003)
- How to deal with a large number of monthly data series?
  - ✓ Factor models (Boivin and Ng, 2005; Forni et al., 2005)

Motivation (cont'd)

Some challenges:

- How to fix the "unbalance" data (Jagged/ragged issue)?
  - ✓ Arbitrary missing observation problem: no observation for some series in the current month; some series released recently may contain missing values at the beginning.
  - ✓ In real time, some data are released at the beginning of the month, some in the middle, and some at the end.
  - ✓ Giannone et al. (2008): principal component analysis with a modified Kalman Filter
- How to reduce model uncertainty?

Motivation (cont'd)

Model uncertainty

- Breiman (2001): In out-of-sample forecasts, "perturbations" to data (adding or deleting a few data points) may cause a dramatic shift of the optimal predictive model.
- Example:
  1. Model 1:

     $$Y = 2.1 + 3.8X_3 - 0.6X_8 + 83.2X_{12} - 2.1X_{17} + 3.2X_{27}$$

  2. Model 2:

     $$Y = -8.9 + 4.6X_5 + 0.01X_6 + 12.0X_{15} + 17.5X_{21} + 0.2X_{22}$$

Motivation (cont'd)

Model Averaging

- A sensible approach to reducing model uncertainty and structural instability is model averaging.
- MA is a smoothed extension of model selection and may result in increased robustness against misspecification biases of individual forecasts (Hsiao and Wan, 2014).
- Two Main MA Approaches:
    - Bayesian model averaging (BMA): Hoeting et al. (1999)
    - Frequentist model averaging (FMA): mainly aims at selecting optimal weights towards the achievement of some form of optimality.

Motivation (cont'd)

Model Averaging

- FMA includes:
  - Mallows criterion (Hansen, 2007; Wan et al., 2010)
  - Jackknife criterion (Hansen and Racine, 2012)
  - Kullback-Leibler type measures (Zhang et al., 2016)
  - leave-subject-out cross-validation (Gao et al., 2016)
  - Local Jackknife criterion (Sun et al., 2021)
- These methods are designed for selecting optimal combination weights for the conventional data, not for the unbalanced data.

Motivation (cont'd)

This Paper

- Propose a new real-time nowcasting forecast combination with dynamic factor regressions.
- This deletes redundant predictors and selects optimal weights for candidate models, simultaneously.
- It is shown the selected weight achieves the asymptotic optimality and consistency, even all candidate models are misspecified.
- The proposed estimator is consistent and asymptotic normality, if a true model is included in candidate models.
- The proposed method is applied to forecast quarterly GDP with a set of 118 macroeconomic monthly data series, which outperforms other competing methods.

Unbalanced Data Structure

Simple Case

- Let $\mathbf{x}_t = (x_{1,t}, \ldots, x_{4,t})'$ be a $n \times 1$ vector denoting at month t.
- y_k be quarterly data at quarter k.
- $v^*_{q,t}$ contains indexes of n^*_q monthly series that are released at the release date (q, t), $q = 1, \cdots, Q$.
- $n^*_q = ||v^*_{q,t}||$ denoting the number of newly released monthly series right before time t.
- Let $v_{q,t} = \bigcup_{i \leq q} v^*_{i,t}$, and $n_q = \sum_{i \leq q} n^*_i$ is the number of available series at (q, t).

Simple Case

- In this case, $v^*_{1,t} = \{1,2\}$, $v^*_{2,t} = \{3\}$, $v^*_{q,t} = \{4\}$, where $v_{2,t} = \{1,2,3\}$, $v_{3,t} = \{1,2,3,4\}$, $\mathbf{x}_{i \in v_{2,t}} = (x_{1,t}, x_{2,t}, x_{3,t})'$.

Table 1: Dataset available between the first release date and the second release date in the second month of current quarter

k		1			2		$\cdots$		K			$K+1$	
t	1	2	3	4	5	6	$\cdots$	$T-4$	$T-3$	$T-2$	$T-1$	T	$T+1$
$\mathbf{x}_{i,t}$	$x_{1,1}$	$x_{1,2}$	$x_{1,3}$	$x_{1,4}$	$x_{1,5}$	$x_{1,6}$	$\cdots$	$x_{1,T-4}$	$x_{1,T-3}$	$x_{1,T-2}$	$x_{1,T-1}$	$x_{1,T}$	NA
	$x_{2,1}$	$x_{2,2}$	$x_{2,3}$	$x_{2,4}$	$x_{2,5}$	$x_{2,6}$	$\cdots$	$x_{2,T-4}$	$x_{2,T-3}$	$x_{2,T-2}$	$x_{2,T-1}$	$x_{2,T}$	NA
	$x_{3,1}$	$x_{3,2}$	$x_{3,3}$	$x_{3,4}$	$x_{3,5}$	$x_{3,6}$	$\cdots$	$x_{3,T-4}$	$x_{3,T-3}$	$x_{3,T-2}$	$x_{3,T-1}$	NA	NA
	$x_{4,1}$	$x_{4,2}$	$x_{4,3}$	$x_{4,4}$	$x_{4,5}$	$x_{4,6}$	$\cdots$	$x_{4,T-4}$	$x_{4,T-3}$	$x_{4,T-2}$	$x_{4,T-1}$	NA	NA
y_k		y_1			y_2		$\cdots$		y_K			y_{K+1}	
$\mathbf{F}_t$	$\mathbf{F}_1$	$\mathbf{F}_2$	$\mathbf{F}_3$	$\mathbf{F}_4$	$\mathbf{F}_5$	$\mathbf{F}_6$	$\cdots$	$\mathbf{F}_{T-4}$	$\mathbf{F}_{T-3}$	$\mathbf{F}_{T-2}$	$\mathbf{F}_{T-1}$	$\mathbf{F}_T$	

Data Generating Process

- First assume the monthly data series $\mathbf{x}_t$ has r common factors

$$\mathbf{x}_t = \eta + \mathbf{\Lambda}\mathbf{F}_t + \boldsymbol{\epsilon}_t,$$

where $\mathbf{x}_t = (x_{1,t}, \cdots, x_{n,t})'$ is a $n \times 1$ monthly data series at month t, for $t = 1, \cdots, T$, $\mathbf{F}_t = (f_{1t}, \cdots, f_{rt})'$ is $r \times 1$ monthly common factors at month t, $\mathbf{\Lambda}$ is the $n \times r$ factor loading matrix with generic entry λ_{ij}, η is the mean vector, and $\boldsymbol{\epsilon} \sim N(\mathbf{0}, \mathbf{\Omega}_{n \times n})$.

- The dynamics of the common factors follows an $AR(1)$ process

$$\mathbf{F}_t = \mathbf{A}\mathbf{F}_{t-1} + \mathbf{u}_t,$$

where $\mathbf{A}$ is a $r \times r$ matrix, all roots of $\det(\mathbf{I}_r - \mathbf{A}z)$ lie outside the unit circle, and $\mathbf{u}_t \sim WN(\mathbf{0}, \mathbf{I}_r)$.

Data Generating Process

- The data generating process (DGP):

$$y_{k+h} = \mu_k + \varepsilon_{k+h} = \beta_0 + \boldsymbol{\beta}(L)'\mathbf{F}_{3k} + \boldsymbol{\alpha}(L)'y_k + \varepsilon_{k+h},$$

  where $\boldsymbol{\beta}(L)$ is a lag polynomial of order p, and $\boldsymbol{\alpha}(L)$ is a lag polynomial of order s.
- The DGP is rewritten as

$$y_{k+h} = \mu_k + \varepsilon_{k+h} = \mathbf{z}_k'\boldsymbol{\beta} + \varepsilon_{k+h},$$

  where $\mathbf{z}_k = (1, \mathbf{F}_{3k}', \ldots, \mathbf{F}_{3k-p_{max}}', y_{k-1}, \cdots, y_{k-s_{\max}})'$.
- How to find the "right" time lag, lag of factors and number of common factors?

Methodology (cont'd)

Candidate Models

- Consider the mth candidate model $(m = 1, \cdots, M)$

$$y_{k+h} = \mathbf{z}_k^{(m)\prime} \boldsymbol{\beta}^{(m)} + \varepsilon_{k+h}^{(m)},$$

  or in matrix notation

$$\mathbf{y} = \mathbf{Z}^{(m)} \boldsymbol{\beta}^{(m)} + \boldsymbol{\varepsilon}^{(m)},$$

  where $\mathbf{y} = (y_{1+h}, \cdots, y_{K+h})'$, $\mathbf{Z}^{(m)} = (\mathbf{z}_1^{(m)}, \cdots, \mathbf{z}_K^{(m)})'$, $\boldsymbol{\mu} = (\mu_1, \cdots, \mu_K)'$ and $\boldsymbol{\varepsilon}^{(m)} = (\varepsilon_{1+h}^{(m)}, \cdots, \varepsilon_{K+h}^{(m)})'$.
- $\mathbb{E}(\varepsilon_{k+h}^{(m)} | \mathbf{z}_k^{(m)}) \neq 0$ is allowed.

Candidate Models

- A big issue for the factor model is we need to determine the number of factors.
- We use one of Bai and Ng (2002) criteria defined by

$$BIC_3(r) = V(r, \widehat{F}^r) + r\widehat{\sigma}^2(\frac{(n+T-r)ln(nT)}{nT}),$$

where $V(r, \widehat{F}^r) = \min_\Lambda \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}(\mathbf{x}_{it} - \lambda_i^{r'}\widehat{\mathbf{F}}_t^r)^2$ is the average residual variance when r factors are assumed.

Methodology (cont'd)

Estimation

- To deal with the "jagged edge" feature, we follow the GRS method (Ginanone et al., 2008) to combine principal component analysis with a modified Kalman Filter to estimate common factors $\mathbf{F}_k$
  - Step 1: OLS estimation on the principle components, based on the truncated monthly series $\mathbf{x}_k$ at time $T-1$
  - Step 2: use Kalman smoother to recursively compute the expected value of the common factors $\mathbf{F}_k$

Estimation

- The unobservable factors $\mathbf{F}$ are replaced by estimators $\widetilde{\mathbf{F}}^{(m)} = (\widetilde{\mathbf{F}}_1^{(m)}, \cdots, \widetilde{\mathbf{F}}_K^{(m)})'$ and thus $\widetilde{\mathbf{Z}}^{(m)} = (\mathbf{z}_1^{(m)}, \cdots, \mathbf{z}_K^{(m)})'$.
- The least squares estimators for $\beta^{(m)}$ and $\boldsymbol{\mu}$ in the mth candidate model are

$$\hat{\beta}^{(m)} = (\widetilde{\mathbf{Z}}^{(m)'}\widetilde{\mathbf{Z}}^{(m)})^{-1}\widetilde{\mathbf{Z}}^{(m)'}\mathbf{y},$$

and

$$\hat{\boldsymbol{\mu}}^{(m)} = \widetilde{\mathbf{Z}}^{(m)}\widehat{\beta}^{(m)} = \mathbf{P}^{(m)}\mathbf{y},$$

where $\mathbf{P}^{(m)} = \widetilde{\mathbf{Z}}^{(m)}(\widetilde{\mathbf{Z}}^{(m)'}\widetilde{\mathbf{Z}}^{(m)})^{-1}\widetilde{\mathbf{Z}}^{(m)'}$.

Candidate Models

- Let $\mathbf{w} = (w^1, \cdots, w^M)'$ be a combination weight vector, which satisfies

$$\mathcal{H}_K = \left\{ \mathbf{w} \in [0,1]^M : \sum_{m=1}^{M} w^m = 1 \right\}.$$

- Unrestricted weights cannot be regularized
  - This may result in poor sampling performance (Hansen, 2012, Lecture note).

Methodology (cont'd)

Candidate models

- The averaging estimator takes the form

$$\hat{\boldsymbol{\mu}}(\mathbf{w}) = \sum_{m=1}^{M} w^m \hat{\boldsymbol{\mu}}^{(m)}.$$

- The corresponding model averaging estimator of parameter $\boldsymbol{\beta}$ is

$$\widehat{\boldsymbol{\beta}}(\mathbf{w}) = \sum_{m=1}^{M} w^m \boldsymbol{\Pi}^{(m)\prime} \widehat{\boldsymbol{\beta}}^{(m)},$$

where $\boldsymbol{\Pi}^{(m)}$ is a matrix of size $q_m \times q$ mapping $\boldsymbol{\beta}$ to $\boldsymbol{\beta}^{(m)}$.

Leave-h-out FV Estimation

- For auto-correlated and heterogenous errors and one-sided data, we use leave-h-out forward-validation estimator $\widetilde{\boldsymbol{\mu}}^{(m)}$.
- Denote a selection matrix $\phi_k^{(m)} = (\mathbf{I}_k, \mathbf{0}_{k\times(K-k)})$ for $1 \leq k \leq h$, and $\phi_k^{(m)} = (\mathbf{0}_{h\times(k-h)}, \mathbf{I}_h, \mathbf{0}_{h\times(K-k)})$ for $h+1 \leq k \leq K$.
- Denote $\mathbf{y}^{(m)}_{[-(t+h)]}$ and $\widetilde{\mathbf{Z}}^{(m)}_{[-k]}$ as the remaining sets of $\mathbf{y}$ and $\widetilde{\mathbf{Z}}^{(m)}$ after removing $\mathbf{y}^{(m)}_{[t+h]}$ and $\widetilde{\mathbf{Z}}^{(m)}_{[t]}$ respectively.
- The leave-h-out FV estimator for $\boldsymbol{\mu}_k$ in the mth candidate model is

$$\widetilde{\mu}_k^{(m)} = \boldsymbol{\pi}_k^{(m)} \widetilde{\mathbf{Z}}^{(m)}_{[k]} \left(\widetilde{\mathbf{Z}}^{(m)\prime}_{[-k]} \widetilde{\mathbf{Z}}^{(m)}_{[-k]} \right)^{-1} \widetilde{\mathbf{Z}}^{(m)\prime}_{[-k]} \mathbf{y}^{(m)}_{[-(k+h)]}$$

where $\boldsymbol{\pi}_k^{(m)} = (\mathbf{0}_{1\times(k-1)}, 1)$ for $1 \leq k \leq h$, and $\boldsymbol{\pi}_k^{(m)} = (\mathbf{0}_{1\times(h-1)}, 1)$ for $h+1 \leq k \leq K$.

Leave-h-out FV Estimation

- The leave-h-out FVMA estimator for μ is

$$\widetilde{\boldsymbol{\mu}}(\mathbf{w}) = (\widetilde{\mu}_1(\mathbf{w}), \cdots, \widetilde{\mu}_K(\mathbf{w}))',$$

  where $\widetilde{\mu}_k(\mathbf{w}) = \sum_{m=1}^{M} w^m \widetilde{\mu}_k^{(m)}$.
- Squared error loss

$$L(\mathbf{w}) = [\widehat{\boldsymbol{\mu}}(\mathbf{w}) - \boldsymbol{\mu}]'[\widehat{\boldsymbol{\mu}}(\mathbf{w}) - \boldsymbol{\mu}].$$

- However, this is infeasible given the unknown conditional mean $\boldsymbol{\mu}$.

Model Averaging Criterion

- Thus, we consider a feasible leave-h-out FV weight choice criterion with a LASSO penalty is

$$\mathrm{FV}(\mathbf{w}) = [\mathbf{Y} - \widetilde{\boldsymbol{\mu}}(\mathbf{w})]'[\mathbf{Y} - \widetilde{\boldsymbol{\mu}}(\mathbf{w})] + \lambda_K \sum_{j=1}^{q} \widehat{v}_j |\widehat{\beta}_j(\mathbf{w})|,$$

  where λ_K is a nonnegative regularization parameter, $\widehat{v}_j = 1/|\widehat{\beta}_j^{LC}|$, $\widehat{\boldsymbol{\beta}}_j^{LC}$ is a least square estimator for $\boldsymbol{\beta}_j$ in the full model, and q is the number of all regressors.

- Minimizing $FV(\mathbf{w})$ with respect to $\mathbf{w}$, we have

$$\widetilde{\mathbf{w}} = \arg \min_{\mathbf{w} \in \mathcal{H}_K} \mathrm{FV}(\mathbf{w}),$$

  and then the leave-h-out FV model averaging (FVMA) estimator of $\boldsymbol{\mu}$ is $\widehat{\boldsymbol{\mu}}(\widehat{\mathbf{w}})$.

Model Averaging Criterion

- The reason of LASSO penalty for parameter sparsity is that the proposed criteria has the potential sparsity of the weights.

Lemma (Sparsity)

Suppose the optimization problem

$$\min_{\mathbf{w}} F(\mathbf{w}) \equiv \mathbf{w}'\mathbf{\Gamma}\mathbf{w}$$
$$\text{subject to } w^m \geq 0, m = 1, \cdots, M,$$
$$\mathbf{1}_M'\mathbf{w} = 1,$$

where $\mathbf{\Gamma}$ *is a nonnegative definite matrix,* $F(\mathbf{w})$ *is linear quadratic and* $\mathbf{1}_M$ *is an* $M \times 1$ *with all elements being 1. Then,* $\exists A \neq \emptyset$*, s.t., a typical solution of* $\mathbf{w}$ *has* $\widehat{w}^m = 0$ *for* $m \in A$.

Asymptotic Theory

Theorem (Asymptotic Optimality)

Suppose Conditions (C.1)-(C.5) hold. The FVMA estimator satisfies the asymptotic optimality (OPT) property, i.e.,

$$\frac{L(\widehat{\mathbf{w}})}{\inf_{\mathbf{w}\in\mathcal{H}_K} L(\mathbf{w})} \xrightarrow{p} 1, \textit{as } K \to \infty.$$

- Theorem I shows that the local squared error loss obtained from the weight vector $\widehat{\mathbf{w}}$ is asymptotically equivalent to the infeasible optimal combination weight vector.

Asymptotic Theory (cont'd)

Consistency

- The model averaging estimator $\widehat{\beta}(\mathbf{w}) = \sum_{m=1}^{M} w^m \mathbf{\Pi}^{(m)\prime} \widehat{\beta}^{(m)}$.
- Define the optimal weight

$$\mathbf{w}^0 = \arg \min_{\mathbf{w} \in \mathcal{H}_K} \mathbb{E}L(\mathbf{w}),$$

and the expected local quadratic loss

$$\widetilde{\xi} = \min_{\mathbf{w} \in \mathcal{H}_K} \mathbb{E}L(\mathbf{w}),$$

where $\mathbf{w}^0 \equiv (w_1^0, \cdots, w_M^0)'$. Let $\zeta_{\min}(\mathbf{A})$ and $\zeta_{\max}(\mathbf{A})$ be the minimum and maximum singular values of matrix $\mathbf{A}$, respectively.

Consistency

Theorem (Consistency of Weights Estimation)

If Conditions (C.1)-(C.10) hold, there exists a local minimizer $\widehat{\mathbf{w}}$ of $FV(\mathbf{w})$ such that

$$||\widehat{\mathbf{w}} - \mathbf{w}^0|| = O_p(K^{-1/2+\delta}\widetilde{\xi}^{1/2}), \text{ as } T \to \infty,$$

where δ is a positive constant given in Condition (C.9).

- Theorem II shows that the estimated combination weight vector $\widehat{\mathbf{w}}$ converges to the optimal weight $\mathbf{w}^0$ in probability at the rate of $\widetilde{\xi}^{1/2}K^{-1/2+\delta}$. The slower the rate of $\widetilde{\xi} \to \infty$, the faster the rate of $\widehat{\mathbf{w}}$ converges to $\mathbf{w}^0$ as $K \to \infty$.

Consistency

Theorem (Consistency of MA Parameter Estimation)

If Conditions (C.1)-(C.10) hold. Then

$$\widehat{\beta}(\widehat{\mathbf{w}}) - \beta^*(\mathbf{w}^0) = O_p(\widetilde{\xi}^{1/2} K^{-1/2+\delta} + K^{-1/2}).$$

where δ is a positive constant given in Condition (C.9).

- Denote $\beta^*(\mathbf{w}^0) = \sum_{m=1}^{M} w_m^0 \beta^{(m)*}$ with $\beta^{(m)*} = \arg\min_{\beta^{(m)}} \mathbb{E}(y_{k+h} - \widetilde{\mathbf{Z}}_k^{(m)'} \beta^{(m)})^2 = [\mathbb{E}(\widetilde{\mathbf{Z}}_k^{(m)} \widetilde{\mathbf{Z}}_k^{(m)'})]^{-1} \mathbb{E}(\widetilde{\mathbf{Z}}_k^{(m)} y_{k+h})$.
- Theorem III indicates that the model averaging estimator $\widehat{\beta}(\widehat{\mathbf{w}})$ converges to a well-defined limit $\beta^*(\mathbf{w}^0)$, even if all candidate models may be misspecified.

Asymptotic Theory (cont'd)

Asymptotic Properties When There Are Correct Candidate Models

- True model: $Y_{k+h} = \mathbf{X}_k' \boldsymbol{\beta}_k + \varepsilon_{k+h}$, where $\mathbf{X}_k$ is a $q \times 1$ vector of explanatory variables, $\boldsymbol{\beta}_k$ is a $q \times 1$ vector and $\mathbb{E}(\varepsilon_{k+h}|\mathbf{X}_k) = 0$ a.s. Define the $(M_0 + 1)$th model is the true model.
- Under-fitted model: Any candidate model omitting regressors with nonzero parameters
- Over-fitted model: Any candidate model which has no omitted variable but has irrelevant variables

Asymptotic Theory (cont'd)

Asymptotic Normality

Theorem (Asymptotic Normality of MA Parameter Estimation)

Suppose Conditions (C.1)-(C.10) hold. Then we have $||\widehat{w}^m|| = O_p(K^{-1/2})$ *for* $1 \leq m \leq M_0$, $\widehat{w}^m = O_p(\lambda_K^{-1})$ *for* $m \geq M_0 + 2$, *and thus* $\sqrt{K}[\widehat{\beta}(\widehat{\mathbf{w}}) - \beta] \xrightarrow{d} N(\mathbf{0}, \mathbf{\Omega}^{(M_0+1)})$, *as* $T \to \infty$.

- It is shown that the selected weights for under-fitted models converge to zero with a faster rate than that for over-fitted models.

Simulation

Data generating process

- The approximate factor models is

$$\begin{aligned} X_{it} &= \lambda_i F_t + \sqrt{r} e_{it}, \\ F_{jt} &= \alpha_j F_{jt-1} + u_{jt}, \\ e_{it} &= \rho_i e_{it-1} + \epsilon_{it}, \end{aligned}$$

  where F_{jt} is the jth component of F_t, $r = 3$, $\lambda_i \sim N(0, rI_r)$, $\alpha_j \sim U[0.2, 0.8]$, $\rho_i \sim U[0.3, 0.8]$, $(u_{jt}, \epsilon_{it}) \sim N(0, I_2)$.

- The predictive regression model is

$$\begin{aligned} y_{k+h} &= y_k + F'_{3k}\beta_1 + F'_{3k-1}\beta_2 + \varepsilon_{k+h}, \\ \varepsilon_{k+h} &= \sum\nolimits_{j=1}^{h-1} \pi^j v_{k+h-j}, \end{aligned}$$

  where $v_k \sim N(0,1)$, $\beta = c(\beta'_1, \beta'_2)' = (0.5, 0.5, 0.2, 0.2, 0.1, 0.1)'$, c is a scaling parameter ranging from 0.2 to 1.2 for $h = 1$.

- For multi-step forecasting, the moving average parameter π ranges from 0.1 to 0.9 and the scale c is held at 1.

Simulation

Table 1. Simulation results: MSFE comparisons

Scale	OPT	GRS	MMA	JMA	SBIC	SAIC
			h=1, lead=2			
0.2	1.0067	1.1426	1.0424	1.0703	1.0301	1.0943
0.3	1.0482	1.2187	1.0947	1.1236	1.0866	1.1506
0.4	1.1048	1.3213	1.1668	1.1968	1.1628	1.2279
0.5	1.1765	1.4461	1.2599	1.2910	1.2586	1.3263
0.6	1.2633	1.5996	1.3738	1.4063	1.3744	1.4457
0.7	1.3652	1.7821	1.5074	1.5423	1.5100	1.5861
0.8	1.4822	1.9919	1.6609	1.6999	1.6656	1.7478
0.9	1.6143	2.2297	1.8345	1.8778	1.8409	1.9310
1	1.7615	2.4949	2.0286	2.0758	2.0359	2.1359
1.1	1.9238	2.7879	2.2423	2.2942	2.2506	2.3624
1.2	2.1012	3.1116	2.4762	2.5341	2.4850	2.6103

Table 2. Simulation results: MSFE comparisons

Scale	OPT	GRS	MMA	JMA	SBIC	SAIC
h=2, lead=2						
0.1	1.7512	2.4333	1.8616	1.8980	1.8385	1.9913
0.2	1.8490	2.5807	1.9033	1.9506	1.9038	2.0418
0.3	1.9522	2.6347	2.0383	2.0784	2.0198	2.1584
0.4	1.9040	2.7258	1.9487	2.0078	1.9872	2.0773
0.5	2.0890	2.8471	2.1828	2.2381	2.1887	2.3244
0.6	2.1957	3.0513	2.2524	2.3118	2.2630	2.3897
0.7	2.3042	3.2644	2.3740	2.4282	2.3545	2.5125
0.8	2.3707	3.3203	2.5759	2.6296	2.5091	2.7894
0.9	2.5204	3.3640	2.6457	2.7131	2.6330	2.8409

Empirical Application

- The monthly data set consists of 118 macroeconomic indictors for Chinese economy collected from 1998 Jan to 2017 June.
- These indicators are divided into 16 blocks and released on 8 dates throughout the months.

Release Date	Block	Number of Series	Total No. of Series
1^{st}	Interest Rate	6	30
	Stock Market	6	
	Foreign Financial	18	
12^{th}	Price Index	5	13
	External One	2	
	Building	5	
	Government	1	
17^{th}	Investment	4	4
19^{th}	Exchange Rate	9	17
	Producer Price Index	8	
21^{st}	World Commodity Price	9	9
24^{th}	External 2	18	19
	Retail Sales	1	
25^{th}	Economic Climate Indicator	3	18
	Industrial Production	15	
28^{th}	Monetary	8	8
Total	16 Blocks	118 Series	

Empirical Application

- The target predictor: China's quarterly GDP growth rate (relative to the nominal GDP of the same quarter of the previous year)
- Cheng and Hansen (2015): MMA and JMA

Table 3. Out-of-sample RMSFE (%) during 2010Q1-2017Q2

horizon	OPT	GRS	MMA	JMA	SBIC	SAIC
h=1, lead=3	1.83	2.28	6.88	1.86	1.78	8.42
h=1, lead=2	1.65	2.15	6.88	1.86	1.78	8.42
h=1, lead=1	1.60	2.00	6.88	1.86	1.78	8.42
h=0, lead=3	1.06	1.75	1.58	1.67	1.60	1.57
h=0, lead=2	0.96	1.61	1.58	1.67	1.60	1.57
h=0, lead=1	0.91	1.47	1.58	1.67	1.60	1.57

Empirical Application

- Within-quarter data flows matter in the sense that the precision of the nowcast generally increases monotonically as new information becomes available during the current quarter.
- Exploiting rich data sets is relevant for real-time data analysis.
- As more information becomes available throughout the quarter, uncertainty declines.

Conclusion

- We have contributed to the forecasting literature by proposing a novel real-time forecasting combination method.
- The asymptotic optimality and consistency of the model averaging estimator are derived.
- The convergence rate of the weights is obtained.
- If the true model is included, the asymptotic normality of the model averaging estimator is derived.
- The proposed method is applied to China's GDP forecast, which compares favorably to the popular nowcasting methods in the existing literature.

Thank you very much!