On the Predictability of China Macro Indicator with Carbon Emissions Trading

Chen, Qian ; Gao, Xiang ; Xie, Shan ; Sun, Li ; Tian, Shuairu ; Hamori, Shigeyuki

On the Predictability of China Macro Indicator with Carbon Emissions Trading

Qian Chen [1], Xiang Gao [2], Shan Xie [3], Li Sun [2], Shuairu Tian [2] and Shigeyuki Hamori [4,*]

[1] Business School, Shenzhen Technology University, 3002 Lantian Road, Pingshan District, Shenzhen 518118, China; chenqian@sztu.edu.cn

[2] Research Center of Finance, Shanghai Business School, 2271 West Zhongshan Road, Shanghai 200235, China; gaoxiang@sbs.edu.cn (X.G.); sunli@sbs.edu.cn (L.S.); tiansr@sbs.edu.cn (S.T.)

[3] Yangtze Ecology and Environment Co., Ltd., Liuhe Road, Jiangan District, Wuhan 430014, China; 1701213123@sz.pku.edu.cn

[4] Graduate School of Economics, Kobe University, 2-1, Rokkodai, Nada-Ku, Kobe 657-8501, Japan

* Correspondence: hamori@econ.kobe-u.ac.jp

Abstract: Accurate and timely macro forecasting requires new and powerful predictors. Carbon emissions data with high trading frequency and short releasing lag could play such a role under the framework of mixed data sampling regression techniques. This paper explores the China case in this regard. We find that our multiple autoregressive distributed lag model with mixed data sampling method setup outperforms either the auto-regressive or autoregressive distributed lag benchmark in both in-sample and out-of-sample nowcasting for not only the monthly changes of the purchasing managers' index in China but also the Chinese quarterly GDP growth. Moreover, it is demonstrated that such capability operates better in nowcasting than h-step ahead forecasting, and remains prominent even after we account for commonly-used macroeconomic predictive factors. The underlying mechanism lies in the critical connection between the demand for carbon emission in excess of the expected quota and the production expansion decision of manufacturers.

Keywords: high-frequency carbon emissions trading; macroeconomic forecast; mixed data sampling regression; GDP growth; purchasing managers' index; out-of-sample prediction

JEL Classification: C53; E27; Q47; Q59

1. Introduction

Predicting macro indicators has always been a hot area of research. This is because accurate and timely forecast serves as an important reference point for economic decision-making by both policymakers and investors. Giannone et al. [1] find that the timeliness, as well as the quality of information, are helpful to increase prediction accuracy when generating nowcasts of quarterly GDP data using current-quarter data when they are immediately available. To deal with the fact that macroeconomic time series are often of different frequencies, researchers resort to the mixed data sampling (MIDAS) regressions. When higher frequency data is evaluated at its original format rather than an aggregated one, the information contained in various data sampled across different frequencies could be retained to the largest extent. What is more, while in traditional co-frequency models we need to wait for the release of data over a full reporting cycle before making predictions, in the MIDAS model one can immediately update the prediction output upon continuous changes in high-frequency predictors. Quite a few studies have applied this MIDAS approach to macro forecasting in China. Though more frequent than the prediction target variables, their chosen predictors are still traditional macroeconomic variables that are limited by some time lag in data publication [2]. Therefore, finding our true high-frequency predictors without any lag, preferably real-time trading data, would help improve the accuracy and timeliness of Chinese macro forecasts.

Given the vital connections between GDP, trade, foreign direct investment (FDI), and carbon emissions in China [3], the goal of the current research is to explore whether carbon emissions trading data could serve as a new powerful predictor for Chinese macro forecasts. Carbon emissions trading is a tool designed to reduce emissions via market power. To implement this tool, the government sets a cap to control the total quantity of carbon emission, and then allocates quota to polluting firms according to a set of rules. The reason why we propose such a predictor is as follows.

The demand for and hence the trading activities of carbon emissions come from the difference between carbon allowances and actual emissions generated in the production process. In general, during economic booms, emissions will increase given that firms expand production and subsequently want more quota to emit carbon. If extra permits are not granted by the government, these demands will be reflected in the trading of carbon emissions. The extant literature focuses on how to price the permit, overlooking the information discovery function. We investigate the predictability of macro indicators by carbon emissions trading by highlighting its informational relevance with the state of the economy and its data availability on a daily basis.

We choose China as a case study for two reasons. One, the task of forecasting Chinese growth has become challenging in recent years after the country's high speed growth phase. Two, the Chinese carbon emissions trading market is highly regulated, reflecting actual demands rather than speculations. The Chinese high-frequency carbon emissions trading data has been available since the establishment of the pilot market in Shenzhen since 2013. And the data quality will continue to improve along with the official launch of the national carbon emissions trading system of China in 2020.

Our empirical strategy is to exploit the information contained in weekly and monthly data on the average price and volume of carbon emissions trading. We treat these two series as high-frequency indicators in both the single-predictor and multiple autoregressive distributed lag model with mixed data sampling method (ADL-MIDAS) framework. The primary output is forecast for the quarterly GDP growth rate in China. Since quarterly GDP has relatively few data points in the relevant sample period, to triple the number of output observations, the monthly manufacturing Purchasing Managers' Index (PMI) growth rate serves as our secondary prediction target. We conduct three comparison analyses. First of all, we investigate whether our ADL-MIDAS model with high-frequency carbon emissions trading data as inputs can outperform a benchmark auto-regressive (AR) model without such data. Second, we move to investigate whether our ADL-MIDAS model can produce superior performance than another benchmark autoregressive distributed lag (ADL) model that employs the same high-frequency carbon emissions trading data. Third, we adopt a combined forecasting model and investigate whether adding high-frequency carbon emissions trading data into the combination can perform well in comparison to the case of using only usual macro forecast factors.

The main results of using carbon emissions trades to forecast Chinese GDP growth are summarized below. When comparing the ADL-MIDAS model with the AR and ADL benchmarks, even using the same high-frequency carbon emissions trading data, our single-predictor and multiple ADL-MIDAS regressions are shown to improve both in-sample and out-of-sample nowcasting performance significantly (i.e., one can think of nowcasting as forecasting in very short intervals, or equivalently in our set up, using carbon emissions trading data close to the quarterly GDP figure release day for the purpose of predicting next-quarter GDP growth rates). This finding indicates that carbon emissions trading data indeed add value, and the MIDAS setup will extract such value by retaining effective information contained in high-frequency indicators for the macro forecast. Then, when comparing the nowcasting results with the h-step ahead forecasts of MIDAS, we conclude that the MIDAS model performs better in terms of nowcasting than forecasting. The reason why nowcasting stands out is that the high-frequency data encompasses too much information which may turn out to be outdated in predicting further away macroeconomic variables. Thus, the advantage of the MIDAS model lies in its timely

forecast. Next, in a combined forecasting setting, our finding is that the combination including high-frequency carbon emissions trading data can generate more accurate and stable out-of-sample forecasts than the combination consisting of only commonly used monthly macroeconomic predictors. In other words, the high-frequency data indeed brings in new messages uncovered by traditional factors. Finally, after we replace the forecast target from GDP to PMI, similar results are obtained—the multiple ADL-MIDAS model incorporating the high-frequency carbon emissions trading data can improve the out-of-sample prediction accuracy for monthly manufacturing PMI growth in comparison to AR and ADL models. Overall, the above-mentioned results imply that the MIDAS model with high-frequency carbon emissions trading data can play an important role in Chinese macro forecasts.

The rest of the paper is organized as follows: Section 2 reviews the background of China's carbon emissions trading market, and the relevant literature on the MIDAS methodology and its applications in a setting similar to ours. Section 3 formally introduces our MIDAS model specification, the benchmarks, as well as the evaluation criteria with respect to prediction accuracy. Section 4 presents the summary statistics of data used, followed by Section 5 that describes the result of how high-frequency carbon emission trading data perform when foreshadowing GDP and PMI growth in China. The last Section 6 concludes or article.

2. Institutional Background and Literature Review

In this section, we first introduce the development of China's carbon emissions trading market. Then, we review the relevant literature on Chinese macro forecasts and our empirical methodology.

The Chinese market for trading carbon emissions emerged along with a global wave of using market mechanisms to mitigate environmental protection issues. To mitigate the greenhouse effect and promote sustainable economic development, countries around the world have reached a series of cooperative agreements such as the Kyoto Protocol. In light of these agreements, the European Union Emissions Trading System (EU ETS) had been officially put into operation in 2005. Since then, to fulfill their respective emissions reduction commitments, protocol participant countries have started to establish their own carbon emissions trading markets. As one of the developing countries, although China has not yet entered the stage of mandatory emissions reduction, it actively participates in the global emissions reduction plan. Specifically, in June 2013, the first regional carbon emissions trading pilot market was initiated in Shenzhen. Subsequently, other markets were established in a number of cities and provinces including Guangdong, Shanghai, Beijing, Hubei, Tianjin, Chongqing, and Fujian.

Carbon emissions trading serves as a tool to reduce emissions through market power. This arrangement functions via properly pricing carbon emissions, providing an incentive for firms to internalize the externality of pollution. Essentially, given binding constraints, companies choose either to cut down production activities to reduce emissions or to pay for carbon emissions beyond the limit. The government sets a cap to control the total quantity of carbon emissions and allocates allowances to each firm according to a set of agreed-upon standards. If a firm needs more carbon emissions quota or has unused quotas, it can buy or sell in the market.

The allocation process of the total carbon allowance is crucial for designing the carbon emissions trading system. There exist two types of allocation methods: free allocation and paid allocation. As for the free allocation, carbon emissions allowances are allocated to polluting enterprises based on a joint consideration of their historical emissions or outputs and some corresponding benchmark levels. In regard to the paid allocation, carbon emissions allowances are distributed through auctions hosted by the government or sold by the government at a fixed price. At the early stage of China carrying out its pilot trading markets, the free allocation method is implemented first. As the carbon emissions trading

market gradually matures in China, the paid allocation method starts to rise in proportion to the free one.

Besides the allocation process, measures taken by the government to ensure firm compliance is another important issue. If a firm's actual carbon pollutants emitted during the inspection period exceed the permitted amount allocated to it, the environmental court will penalize the firm for breaking government regulations. The existing penalties include issuing a rectification order with a time limit, recording misbehaviors in a national firm credit system, reporting to public administrators and tax authorities, blacklisting carbon emission violators, canceling various rights such as previously granted tax exemptions, qualification for winning awards, and future subsidies. This punishment system for non-compliant manufacturers guarantees that heavy polluting firms will purchase extra carbon emission allowances since the overall cost of just polluting beyond the limit is extremely high.

In 2019, the annual allowances transacted in Chinese carbon emissions trading pilot markets have reached an amount of 93 million tons and a turnover of 1.57 billion Yuan. Taking the Guangdong market as an example, which has the largest transaction amount among all pilot markets in 2019, trading companies come from six industries including civil aviation, steel, cement, paper, petrochemical, and power generation. In each industry, the criterion for an individual company being qualified to trade is emitting 20,000 tons of carbon dioxide or consuming energy that is equivalent to 10,000 tons of standard coal per year. According to the Department of Ecology and Environment of Guangdong Province, free quote accounts for a proportion of 95% to 97% of total emissions depending on which industry the firm belongs to.

Early studies on carbon emissions trading mainly focus on developing theories about price formation. With the operation of EU ETS and associated trading data, there is a growing strand of empirical research about testing theoretical pricing mechanisms as well as the volatility and spillover effect of carbon emissions trading prices on other markets or even the entire economy. Turning to the China case, due to a short history of Chinese carbon emissions trading, the extant research on such markets in China is still in its infancy. While most scholars investigate potential determinants of carbon emissions trading price in China, such as the weather, energy price, government policies, financial market fluctuations, and the macro-economy [4,5], others have discussed volatility spillovers of carbon trading prices. However, to our knowledge, no one has utilized information in this price to conduct macro forecasts in China. Our paper aims to fill this gap.

Now, let us focus on two lines of literature that are closely related to the present study: one that links carbon emissions trading to the operation of the macro-economy, and the other that highlights the application of MIDAS regressions in the macro forecast.

Using data published by the EU ETS, many empirical papers have attempted to attribute a large part of the pricing formation of carbon emissions to macroeconomic factors. e.g., Alberola et al. [6] considered the fluctuation in economic activity as a key driving force of carbon price changes in Germany, the U.K., Poland, and Spain. Chevallier [7] finds that the economic conditions in Europe influence prices of EU ETS carbon emissions with a lag, and this is due to institutional constraints. Zou and Wei [8] demonstrated that the spot prices of certified carbon emissions (CERs) traded on the BlueNext exchange are significantly and positively affected by macroeconomic status proxied by the EU-27 index of industrial production index. With respect to China's carbon emissions trading, there are relatively fewer relevant studies. Zhou and Xu [9] employed VAR and VEC models to explore the influencing factors of Chinese carbon trading prices, and the results show that the CSI Industrial Index is positively associated with prices of carbon emissions. In addition, Zeng et al. [5] resorted to a structural VAR model to study the same question. Statistically significant and positive correlations are found between economic growth and carbon trading prices in China.

Given that the existing literature focuses on how carbon trading prices are determined by the macroeconomic conditions, it is surprising that few works are conducted to examine

the informational contents of these pricing and traded quantity data. Their macro forecast implication is a possible direction to explore as long as proper econometric tools are adopted to reconcile the frequency mismatch between high-frequency carbon emissions trading proxies with quarterly macroeconomic indicators. One potential solution is the MIDAS model, which can be regarded as an extension of the ADL. Unlike traditional co-frequency models that often require all the input variables to be standardized at the same frequency, the most prominent advantage of the MIDAS framework is that it can directly deal with data with different frequencies. Therefore, by using MIDAS, we could avoid problems such as losing high-frequency information during aggregation or obtaining false data points when the interpolation method is used to transfer low-frequency variables to high-frequency ones.

The analytical framework of MIDAS is developed by Ghysels et al. [10–12]. And it is originally applied to test risk-return trade-offs and predict financial market volatility. Following studies continue to explore the time dimension of the MIDAS method [13,14]. They propose the time-series linear regressions of the MIDAS model and discuss various augmented versions of MIDAS such as adding weighting schemes and ordered lags into the plain MIDAS model. Andreou et al. [15] first derive the asymptotic properties of the MIDAS estimator, and then use Monte-Carlo simulations to prove that this estimator is actually more efficient than those estimated from traditional co-frequency models in which an equal weighting scheme is adopted and constraints on the underlying data generating process are imposed. The MIDAS models are also extended to combine with other models. For example, Marcellino and Schumacher [16] introduce FA-MIDAS as a joint model of the MIDAS and the factor model. Breitung and Roling [17] propose a non-parametric version of the standard MIDAS model. Foroni et al. [18] develop an unrestricted MIDAS model (U-MIDAS), which allows for the case of no weighting schemes. Barsoum and Stankiewicz [19] incorporate Markov transformation into the U-MIDAS model to forecast GDP growth over different phases of business cycles.

The advantage of MIDAS in dealing with data at different frequencies has been frequently explored in macro forecasts. Marcellino and Schumacher [16] work in their FA-MIDAS model to predict German GDP. Clements and Galvão [20] rely on monthly indicators to foreshadow the U.S. quarterly GDP growth. Clearly, the MIDAS model exhibits advantages in short-term GDP forecasts. Ghysels and Wright [21] integrate the MIDAS model and the Kalman filter so that daily data can be used to predict the GDP growth, inflation rate, unemployment rate, etc. Clements and Galvão [22] employ a multiple MIDAS with a list of leading indices to predict U.S. GDP growth. They show that the MIDAS model performs better than a benchmark AR model. Kuzin et al. [23] use both the MIDAS and M-VAR model to predict European GDP; their result is that: the MIDAS model outperforms in the short run, while the M-VAR model is superior in the long run. Bai et al. [24] compare the performance of the MIDAS model and the State Space model on predicting the U.S. quarterly GDP growth. The MIDAS model again excels in terms of timeliness. Andreou et al. [25] employ nearly 1000 different daily financial indicators to predict the U.S. quarterly GDP growth. Indeed, high-frequency financial data plays a role in the macro forecast. Bessec and Bouabdallah [26] construct an FA-MIDAS model with Markov transformation, in which they also utilize high-frequency financial data to predict U.S. GDP growth. They obtain good in-sample estimates in their analytical framework.

Our paper follows these empirical works by further investigating how a particular type of high-frequency data, carbon emissions trading proxies which combine the characteristics of high-frequency financial data and the activities of manufacturers, performs in forecasting macroeconomic indicators by using China as a case study. A few papers also apply MIDAS in Chinese macro forecast, confirming the superior performance of the MIDAS family with high-frequency inputs in predicting contemporaneous China's quarterly GDP growth. Yet no research has explored the relative importance of carbon emissions trading markets. We step into this blank.

3. Empirical Methodology

The MIDAS regression model is designed to deal with data of different frequencies in the same modeling framework. Let us start with the fundamentals of single-predictor and multiple MIDAS models.

3.1. Basic Setup

The single-predictor MIDAS (m, K) is the most basic form of all, where m is the frequency ratio calculated by comparing the high-frequency independent variable and low-frequency dependent variable, and K denotes the maximum order of lags in the independent variable. For example, if the dependent variable y is sampled quarterly and the independent variable x is a monthly indicator, then we have 1 observation for y and 3 observations for x every quarter. Hence, the frequency ratio m in this example equals 3. The regression specification is given by:

$$y_t = \mu + \beta W(L; \boldsymbol{\theta}) x_{tm} + \epsilon_t, \tag{1}$$

where the weighting function $W(L; \boldsymbol{\theta}) = \sum_{k=0}^{K} w(k; \boldsymbol{\theta}) L^k$. In this definition, L^k represents the lag operator of the high-frequency indicator x such that $L^k x_{tm} = x_{tm-k}$, and $w(k; \boldsymbol{\theta})$ describes the weighing scheme with parameter vector $\boldsymbol{\theta} = (\theta_0, \theta_1, \ldots, \theta_p)$. The non-linear least squares (NLS) approach is usually used to estimate the parameters in Equation (1).

If the restriction on the weighting function is removed, then an unrestricted MIDAS model (U-MIDAS) can be obtained with the following form [18]:

$$y_t = \mu + \sum_{k=0}^{K} \beta_k x_{tm-k} + \epsilon_t. \tag{2}$$

The ordinary least squares (OLS) methodology can be used to estimate the vector of coefficients $\boldsymbol{\beta} = (\beta_0, \beta_1, \ldots, \beta_K)$. To have a better understanding, we rewrite Equation (2) in the matrix notation. Suppose we use a monthly predictor x to forecast a quarterly target variable y given a frequency ratio $m = 3$. By assuming that monthly data in the current and the previous two quarters all have explanatory power on y, the MIDAS model characterizing this example can therefore be expressed as:

$$\begin{bmatrix} y_3 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_9 & \cdots & x_1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{3n} & \cdots & x_{3n-8} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \vdots \\ \beta_9 \end{bmatrix} + \begin{bmatrix} \epsilon_3 \\ \vdots \\ \epsilon_n \end{bmatrix}. \tag{3}$$

There are ten parameters to be estimated in this specification.

Now we move on to introducing multiple high-frequency independent variables, say, n variables. The specification of the multiple MIDAS model (M-MIDAS) is:

$$y_t = \mu + \sum_{i=1}^{n} \beta_i W^{(i)} \left(L; \boldsymbol{\theta}^{(i)} \right) x_{tm_i}^{(i)} + \epsilon_t, \tag{4}$$

where m_i is the frequency ratio of the i-th high-frequency independent variable. Of course, m_i can vary across different independent variables. For a quarterly dependent variable, it is possible to include in Equation (4) both monthly ($m = 3$) and weekly ($m = 12$) independent variables. Finally, we utilize in the empirical section of this paper the version adding a list of autoregressive terms of the dependent variable on the right-hand side. The ADL-MIDAS model with a single predictor is given by:

$$y_t = \mu + \sum_{j=1}^{p_y} \alpha_j y_{t-j} + \beta W(L; \boldsymbol{\theta}) x_{tm} + \epsilon_t, \tag{5}$$

where p_y is the maximum lagged order of y. This treatment considers the persistent effects in the dependent variable as the predictive target.

3.2. Weighting Scheme and Estimation Method

Comparing to traditional models using input data of the same frequency, such as the ADL model, the MIDAS model takes more advantage of the high-frequency data via implementing flexible weighting schemes. This section discusses advantages and problems with the unrestricted and restricted weights respectively.

First, unrestricted MIDAS can be estimated via OLS and is straightforward to interpret. It can retain as many lag periods of the high-frequency predictor (K as large) as the sample size is large enough to ensure effective estimation. However, the U-MIDAS model has a parameter proliferation issue. That is, the number of parameters in U-MIDAS increases with the number of predictors, data frequency of predictors, as well as the number of lags included in the model. U-MIDAS is only applicable if the frequency difference between the predictors and the target variable is not large, that is when the frequency ratio m is small.

On the other hand, a weighting function can be imposed in MIDAS to address the parameter proliferation of U-MIDAS and add parsimony to the model. Ghysels et al. [27] illustrate that, when the data generating process follows a certain constraint, the MIDAS estimator under a well-tuned weighting scheme turns out to be more efficient than the U-MIDAS one. Even an incorrect constraint might be useful in terms of efficiency as long as the sample size of the data under consideration is small.

Therefore, the key elements of MIDAS are determining the optimal weighting scheme $w(k; \boldsymbol{\theta})$ and the optimal lag order K. Following the suggestions made on the weighting scheme of MIDAS [14], this study selects three types of weighting schemes commonly used in the literature, namely, the Almon polynomial, exponential Almon polynomial, and Beta distribution polynomial. A great deal of flexibility can already be achieved with a few weighting parameters summarized in $\boldsymbol{\theta}$. Taking the Almon polynomial weighting scheme as an example, two to four parameters is sufficient to achieve a high degree of flexibility, that is, $p = 1$, 2, or 3 in:

$$w(k; \boldsymbol{\theta}) = \theta_0 + \theta_1 k + \theta_2 k^2 + \ldots + \theta_p k^p. \tag{6}$$

The exponential Almon polynomial weighting scheme is obtained by augmenting the Almon polynomial one:

$$w(k; \boldsymbol{\theta}) = \exp\left(\theta_0 + \theta_1 k + \ldots + \theta_p k^p\right) / \sum_{k=1}^{K} exp\left(\theta_0 + \theta_1 k + \ldots + \theta_p k^p\right). \tag{7}$$

Finally, the Beta distribution polynomial weighting scheme is given by:

$$w(k; \boldsymbol{\theta}) = f\left(\frac{k}{K}, \theta_1, \theta_2\right) / \sum_{k=1}^{K} f\left(\frac{k}{K}, \theta_1, \theta_2\right), \tag{8}$$

where $f(x, a, b) = x^{a-1}(1 - x)^{b-1}\Gamma(a + b)/\Gamma(a)\Gamma(b)$, and $\Gamma(a) = \int_0^\infty e^{-x}x^{a-1}dx$.

Figure 1 plots the weights of lag indexes for each weighting scheme. The two columns of Figure 1 represent two different sets of parameter values, corresponding to fast decaying weighting and slow increasing weighting, respectively. For example, the first figure in the left column plots the weights for 50 lags from an exponential Almon polynomial weighting scheme with $\boldsymbol{\theta} = (1, -0.2)$. In time-series prediction, such asymmetric weights are consistent with the intuition that predictor data from a long time ago should have a lesser effect than more recent data on the predictive power for movements in the target variable. However, traditional models with an equal weighting scheme miss this point. This contrast is evident in Andreou et al. [15], where the authors find that the MIDAS estimator is more efficient than those estimated by traditional models with equal weighting scheme, especially when the underlying data generating process follows the constraint that high-

frequency predictors indeed have predictive power for the dependent variable. Another benefit of using a weighting scheme is the convenience of selecting lag orders. In the U-MIDAS model, the lag order of a predictor usually needs to be determined by repeated attempts if without theoretical support. However, if the appropriate weighting scheme is selected, the estimated weight of the lagged terms can approach zero asymptotically as the lag order increases. Thus, one can determine the best lag order by discarding the lagged term whose weights are close to zero [14]. The lag order is simultaneously determined as the weighting parameters are estimated, thus the lag selection in MIDAS is largely driven by data. However, it is possible that a carefully chosen parameterized weighting scheme leads to information loss as the "smoothness" constraint "shrinks" the unrestricted function and reduces sampling noise. As pointed out by Sinko [28], there is a tradeoff between information loss and noise reduction, but the ultimate importance lies in that whether the MIDAS model produces sensible forecasts.
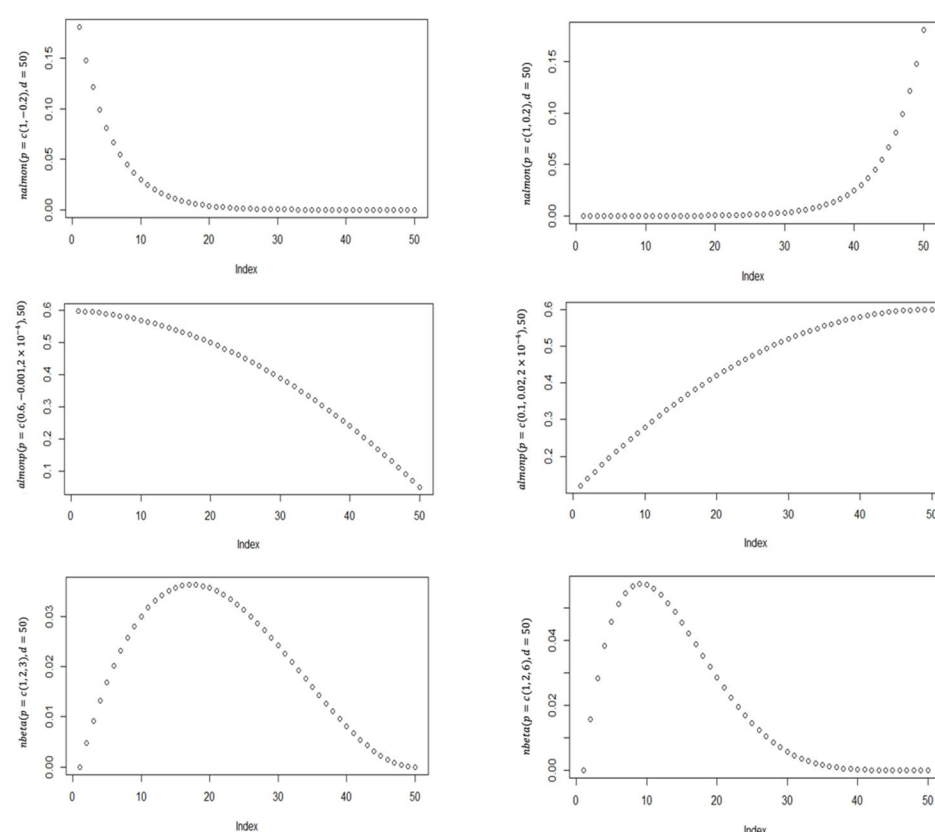
Figure 1. Flexible forms of weighting schemes.

With a weighting scheme and the associated constraints imposed on the MIDAS model, we normally utilize the NLS for parameter estimation. First, an initial set of values is set for the parameters $\boldsymbol{\theta} = (\theta_0, \theta_1, \dots, \theta_p)$ in the weighting scheme. Our model then becomes a linear regression and thus the OLS could be used to estimate the coefficients $\boldsymbol{\beta} = (\beta_0, \beta_1, \dots, \beta_K)$. Given the first-round solution, the final solution can be obtained via implementing an iterative method. One may have a concern that the initial values of the parameters $\boldsymbol{\theta}$ might exert significant influences on the final results. We argue this is not the case by repeating our exercises with a variety of choices on the initial values and by showing that the optimal estimation is independent of these choices. As for determining the optimal weighting scheme of $w(k; \boldsymbol{\theta})$ and the optimal lag order of K, various selection criteria are applied in this research, such as the Akaike information criterion (AIC), Bayesian information criterion (BIC), root-mean-square deviation (RMSE), mean squared forecast errors (MSFE), etc.

3.3. H-Step Ahead Forecast of MIDAS

From now on, we distinguish between nowcasts and forecasts. A nowcast, a.k.a. a real-time forecast, refers to the use of current data to predict the target variables in the same current period. In contrast, an h-step ahead forecast means using current data to predict the target variables in the future. The h-step ahead forecast model of MIDAS can be written as:

$$y_{t+h|t} = \mu + \beta W(L; \boldsymbol{\theta}) x_{tm} + \epsilon_{t+h}. \tag{9}$$

When *h = 0*, the above equation represents a typical nowcast or real-time forecast setup. One prominent advantage of MIDAS is that we can keep updating the high-frequency predictors with the latest data during the period $h = (m-1)/m$ to $h = 0$. This is in contrast to the co-frequency model, in which *h* can only take integer values. For example, the National Bureau of Statistics of China often releases the data on the third-quarter GDP growth in mid-October. Thus, the AR (1) model can only make a prediction when the second-quarter GDP growth is public in mid-July (in this case *h = 1*). But if we use the MIDAS model and monthly predictors, our prediction for the third-quarter GDP growth could be updated every month by taking into account the latest information revealed by the monthly predictors. With weekly or even daily frequency indicators, the macro forecast can be updated more frequently.

3.4. Benchmark Models

We list in this subsection three benchmark models whose results we will compare our MIDAS outputs to. The first benchmark is the AR Model as the autocorrelation is usually presented in the time series of macroeconomic variables due to the inertia of economic systems. The AR model uses historical data of the dependent variable to predict its future value. A *p*-order autoregressive model AR(*p*) is given by:

$$y_t = \mu + \sum_{j=1}^{p} \alpha_j y_{t-j} + \epsilon_t. \tag{10}$$

By setting the AR model as the first benchmark, we aim to evaluate whether our proposed ADL-MIDAS model with high-frequency carbon emission trading data perform better than models without such trading data in macro forecasts.

The second benchmark is the ADL model. Autoregressive distribution lag can represent traditional models requiring variables sampled at the same frequency.

$$y_t = \mu + \sum_{j=1}^{p_y} \alpha_j y_{t-j} + \sum_{i=0}^{q_x} \beta_i x_{t-i} + \epsilon_t. \tag{11}$$

One can view ADL as a simple vision of ADL-MIDAS in Equation (5) with unrestricted weight schemes on the predictors as in Equation (2). By establishing the ADL model as the second benchmark, we would like to know whether our proposed ADL-MIDAS model with high-frequency carbon emission trading data can outperform models that also adopt high-frequency carbon emissions trading data but use specifications different from MIDAS.

Thirdly, we consider a combined forecast model. A combination forecast is defined in the sense that we arrive at multiple sets of prediction values by using multiple models. And the final prediction output is a weighted average of a variety of predictive values. The weights follow a chosen weighting function, e.g., the BIC weighting function is listed as follows:

$$\hat{y}_t = \sum_{k=0}^{K} w_k \hat{y}_{k,t}; \tag{12}$$

$$w_k = e^{-BIC_k} / \sum_{k=1}^{K} e^{-BIC_k}. \tag{13}$$

What is good about the combined forecast model is that it can synthesize the prediction results of different models and overcome potential uncertainty and instability in forecasting models. It is sometimes imperative to consider different models at the same time because certain predictors may require the use of specific models which might not be necessarily compatible under other models' framework. The relevant research shows that combination forecasting might be more stable than forecasts generated by individual models [29–32]. We follow Andreou [33] by employing four kinds of weighting functions in our combined forecast model, i.e., the equal weights (EW), BIC, MSFE, and discounted MSFE (DMSFE).

3.5. Forecasting Accuracy Index

The models are applied to real data and their performance on the out-of-sample forecast is evaluated. A proportion of the sample is held as a forecasting (testing) sample, say one-third or one-fifth of the sample. A fixed window forecasting scheme is adopted in this paper. For example, consider a one-step-ahead forecast with a forecasting sample that is last on third of the whole sample period, denoted 1/3 L. For the first period in the forecasting sample, T+1, where T = 2/3 L, the first T periods data are used to estimate the model, then forecast y_{T+1}. Next, use the fixed window, that is the first T+1 periods' data to re-estimate the model, of course, the weights for each lag are re-estimated, which is equivalent to re-selecting the number of lags. Then forecast y_{T+2}, so on and so forth, and obtain all 1/3 L out-of-sample forecasts and evaluate the prediction errors. The reason for using a fixed window, instead of a rolling window is that the sample size is relatively small due to the short history of the carbon emission markets in China.

The *RMSE* is used in this paper as a basis to measure predicted errors and the relative accuracy of our macro prediction. The formula for calculating *RMSE* is given by:

$$RMSE = \sqrt{\frac{1}{m}\sum_{i=1}^{m}(\hat{y}_i - y_i)}. \tag{14}$$

The ratio of the *RMSE* of our proposed model to the *RMSE* of the first benchmark model AR is denoted by *rRMSE*:

$$rRMSE = \frac{RMSE\ of\ underlying\ model}{RMSE\ of\ AR}. \tag{15}$$

As a result, *rRMSE* < 1 indicates that our model has smaller errors, hence better forecasting performance than AR. Both the mean absolute scaled error (*MASE*) and mean absolute percentage error (*MAPE*) serve as auxiliary evaluation criteria.

$$MAPE = \frac{1}{m}\sum_{i=1}^{m}\left|\frac{\hat{y}_i - y_i}{y_i}\right|; \tag{16}$$

$$MASE = \frac{\frac{1}{m}\sum_{i=1}^{m}|\hat{y}_i - y_i|}{\frac{1}{m-1}\sum_{i=2}^{m}|y_i - y_{i-1}|}. \tag{17}$$

Although a smaller-than-1 *rRMSE* confirms a better forecasting performance of our model, it is often hard to verify whether this outperformance is statistically significant, especially when the statistics of *rRMSE* is close to 1. To deal with this flow, the Diebold-Mariano test is used here, which is a method to compare prediction accuracy and test whether the difference is statistically significant [34]. The null hypothesis is that there is no significant difference in prediction accuracy. Let the *h*-step ahead prediction error at time *t* from model 1 and that from model 2 be denoted by $\varepsilon_{t+h|t}^1$ and $\varepsilon_{t+h|t}^2$.

$$\varepsilon_{t+h|t}^1 = y_{t+h} - y_{t+h|t}^1$$

$$\varepsilon_{t+h|t}^2 = y_{t+h} - y_{t+h|t}^2$$

$$H_0 : \ E\left[L\left(\varepsilon_{t+h|t}^1\right)\right] = E\left[L\left(\varepsilon_{t+h|t}^2\right)\right]$$

$$H_0 : \ E[d_t] = 0, \ d_t = L\left(\varepsilon_{t+h|t}^1\right) - L\left(\varepsilon_{t+h|t}^2\right)$$

$$H_1 : \ E[d_t] > 0, \ one \ sided \ test,$$

where $L\left(\varepsilon_{t+h|t}^1\right)$ is a loss function measuring the accuracy of each forecast (usually in the form of squared error loss or absolute error loss). The Diebold-Mariano test statistic is given by:

$$S = \frac{\bar{d}}{\left(L\hat{R}V_{\bar{d}}/T\right)^{1/2}} \sim N(0,1),$$

and $L\hat{R}V_{\bar{d}}$ is a consistent estimate of the asymptotic long-run variance of $\sqrt{T}\bar{d}$. If the *p*-value of the Diebold-Mariano test statistic falls below 0.05, then we reject the null hypothesis of equal predictive accuracy at the 5% level. We choose to conduct the one-sided test in this study. Consequently, the alternative hypothesis is that our proposed ADL-MIDAS model produces more accurate forecasts than the benchmark model. It is noteworthy that we apply the DM test to the one-step-ahead forecasts only (see Section 5), and the forecast horizon is 1 in the calculation of $LRV_{\bar{d}}$.

4. Data

High-frequency carbon emission trading data is investigated as macro predictors in the present paper, including the monthly and weekly trading prices and volume. We process these trading data by taking the first-order difference of their logarithm values. The target variables include quarterly GDP growth and monthly manufacturing PMI growth in China. In this form, the time series does not suffer from seasonal effects. We control for the predictability of three common macroeconomic factors: the total value of imports and exports, total retail sales of consumer goods, and investment in fixed assets. The time range of the original data spans from October 2013 to December 2019, and the data is sourced from the National Bureau of Statistics and Wind database. All variables in our setup are in growth rate format. Except for the growth rate for carbon emissions trading prices, all other growth rate proxies refer to the real growth rates based on constant price levels. Table 1 lists the definitions and sources of our variables, and Table 2 includes the corresponding summary statistics after processing. The ADF stationary test results tell us that all processed data except total retail sales of consumer goods (*cv2*) are stationary at a significance level of 10%. As the main target variable and predictors are stationary, and *cv2* is used as one control variable in the combined forecast analysis, in order to be consistent in the measurement of the variables (growth rate or percentage changes), *cv2* is retained as it is in this study.

The original data on high-frequency trading of carbon emissions in pilot Chinese markets is available on a daily basis. To obtain comparable variables in all frequencies, we aggregate daily transactions into weekly, monthly, and quarterly data. Although the MIDAS model can directly deal with mixed frequency data, it requires that the frequency ratio between high-frequency predictor and low-frequency target variable remains unchanged. This requirement causes some difficulties in dealing with weekly data since the number of weeks in each month is not fixed over time. The operation manual of the R package *midasr* suggests us to assume that one month has exactly four weeks [27]. So, we need to fix the number of weeks in any given month, taking 4 observations for a month. When a particular month has 5 observations, the first week is deleted. When there are less than 4 observations in a particular month, we use the average of the price from the previous week and that from the next week to fill in the missing trading price. The missing trading volume is hence recorded as 0. There are in total 4 weeks of zero volume data out of 252 observations in the weekly sample of carbon emissions trading.

Table 1. Variable definitions and data sources.

Symbol	Variable Definitions	Frequency	Observations	Data Source
GDP	GDP Growth Rate	Quarterly	21	
PMI	PMI Growth Rate	Monthly	63	
cv1	Investment in Fixed Assets, Accumulated Growth Rate	Monthly	63	National Bureau of Statistics of China
cv2	Total Retail Sales of Consumer Goods, Growth Rate	Monthly	63	
cv3	Total Value of Imports and Exports, Growth Rate	Monthly	63	
mprice	ΔLN(monthly trading price)	Monthly	63	
mamount	ΔLN(monthly trading amount)	Monthly	63	
wprice	ΔLN(weekly trading price)	Weekly	252	Wind
wamount	ΔLN(weekly trading amount)	Weekly	252	

Table 2. Stationarity test and summary statistics.

	ADF Test-Statistics	Lag	Mean	Median	Max	Min	Standard Deviation
GDP	−3.377 *	5	0.0688	0.0690	0.0770	0.0600	0.0043
PMI	−1.91 *	4	−0.0026	−0.0039	0.0517	−0.0494	0.0224
cv1	−1.71 **	5	0.0813	0.0895	0.1974	−0.0028	0.0632
cv2	−2.948	5	0.0900	0.0960	0.1220	0.0490	0.0193
cv3	−1.543 *	5	0.0094	0.0172	0.2043	−0.2102	0.0893
mprice	−3.144 *	4	−0.1223	−0.1429	0.9441	−0.9697	0.3951
mamount	−3.421 **	5	0.5533	0.2874	3.7160	−1.9636	1.1103
wprice	−3.428 **	10	−0.1260	−0.1551	1.3315	−1.1697	0.4602
wamount	−4.124 ***	10	0.5855	0.3672	5.4342	−3.6966	1.5009

Note: The symbols ***, **, and * represent significance under significant levels of 1%, 5%, and 10%, respectively. And all the variables in this table are growth rate data.
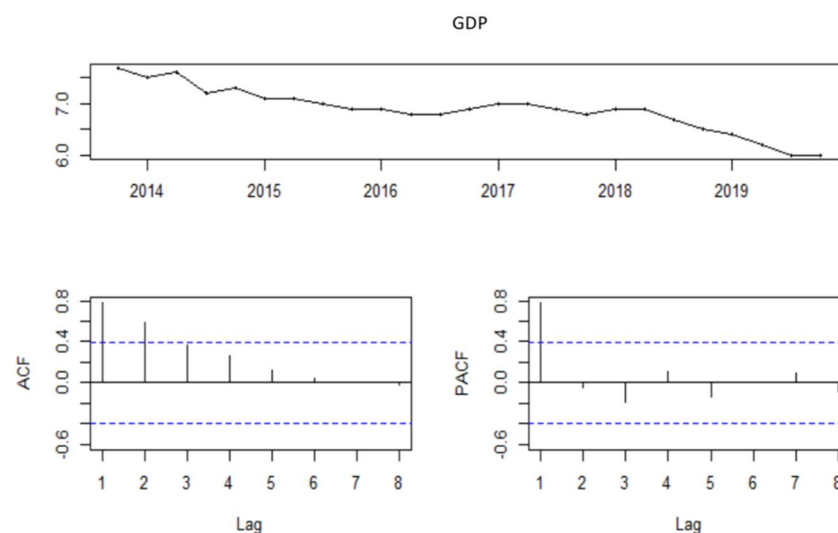
5. Results

5.1. Forecasting Results for Quarterly GDP Growth

For assessing the forecasting performance, the whole dataset is divided into a training set and a test set. The training set is used to estimate models, and the test set is used to calculate the forecasting accuracy index and perform the test. The most recent 1/5 (denoted as 1/5 L) and the most recent 1/3 (denoted as 1/3 L) of the total sample are chosen respectively as the test sets for robustness check.

Both single-predictor and multiple ADL-MIDAS models are considered. The predictors are the trading price change and trading amount change, both have weekly and monthly frequencies. Thus the proposed forecasting models for GDP are (1) ADL-MIDAS with monthly price change (*mprice*); (2) ADL-MIDAS with monthly amount change (*mamount*); (3) multiple ADL-MIDAS with both *mprice* and *mamount*; (4) ADL-MIDAS with weekly price change (*wprice*); (5) ADL-MIDAS with weekly amount change (wprice); (6) multiple ADL-MIDAS with both *wprice* and *wamount*. Model index (1)–(6) will be used in the tables presenting forecasting results.

Considering the short history of the carbon trading pilot market in China, thus a relatively small sample for quarterly GDP in the relevant time horizon, the U-MIDAS model, which means a MIDAS model without weighting scheme, is not considered. The optimal lag order and optimal weighting scheme are determined based on AIC by the model selection function from the R package *midasr* [27]. Two benchmark models are chosen, one is the AR model without any trading data, and the other is the ADL model with the same frequency trading data (quarterly price and amount).

Based on the autocorrelation and partial autocorrelation of GDP growth in Figure 2, we select the AR(1) term for the benchmark AR and ADL models, as well as for ADL-MIDAS models. In practice, we also test AR terms with higher lag orders for the significance of the coefficient and the fitting effect. It turns out that lag order 1 is still the most appropriate and parsimonious choice.

Figure 2. Autocorrelation and partial autocorrelation graph of GDP growth.

Now we are going to discuss the nowcasting of GDP. The forecasting results from the above models are evaluated by *rRMSE*, both in-sample and out-of-sample forecasts, as well as the Diebold-Mariano test for out-of-sample forecast accuracy difference between the proposed MIDAS models and the benchmark AR model. Specifically, we choose the one-sided test, so our alternative hypothesis in this section is that the underlying model is more accurate than the benchmark AR model.

Firstly, according to the *rRMSE* results in Table 3, it can be found that the multiple ADL-MIDAS models consistently have *rRMSE* smaller than 1, and outperform the benchmark AR and ADL model in nowcasting GDP, both in-sample and out-of-sample. Specifically, multiple ADL-MIDAS with weekly predictors generate the smallest *rRMSE* across all models for in-sample prediction with both test sets. For the out-of-sample forecast, ALD-MIDAS with *mamount*, and multiple ALD-MIDAS with monthly predictors generate the smallest *rRMSE*, for the 1/5 L and 1/3 L test set, respectively. As for single-predictor models, they have better in-sample performance than the benchmark AR model, but the out-of-sample prediction performance is less stable. Except for the single-predictor model with *mamount*, the out-of-sample *rRMSE* of the other three single-predictor models at least once exceeded 1, which indicates that they failed to outperform the benchmark model AR in one of the test sets. The ADL model fails to outperform the benchmark AR model for out-of-sample forecasts in both test sets, though the in-sample performance is better than the AR model.

Secondly, according to the Diebold-Mariano test results presented in Table 4, most MIDAS models previously superior to the AR model in the out-of-sample *rRMSE* comparison can reject the null hypothesis at a significant level of at least 10%. For example, for the 1/3 L test set, the *p*-value of the multiple ADL-MIDAS model with monthly trading data is 0.0537, and the *p*-value of the multiple ADL-MIDAS model with weekly trading data is 0.02925, which indicates that this model has significantly better predictive accuracy than benchmark model AR.

Table 3. GDP nowcasting performance comparison by *rRMSE*.

Test Model	Predictors	1/5 L as Test Set		1/3 L as Test Set	
		In-Sample	Out-of-Sample	In-Sample	Out-of-Sample
(1)	*wprice*	0.669	1.408	0.684	1.231
(2)	*wamount*	0.888	0.762	0.667	1.272
(3)	*wprice, wamount*	0.474	0.798	0.587	0.747
(4)	*mprice*	0.665	1.209	0.783	0.673
(5)	*mamount*	0.909	0.588	0.983	0.700
(6)	*mprice, mamount*	0.505	0.756	0.615	0.658
ADL	*price, amount*	0.722	1.359	0.873	1.021
AR		1	1	1	1

Note: The model with the smallest *rRMSE* is highlighted in a box for each time series.

Table 4. GDP nowcasting performance comparison by Diebold-Mariano test.

Test Model	Predictors	H_1: Test Model Is More Accurate than AR Model	
		1/5 L as Test Set	1/3 L as Test Set
(1)	*wprice*	*p*-value = 0.9569	*p*-value = 0.9918
(2)	*wamount*	*p*-value = 0.02901	*p*-value = 0.9273
(3)	*wprice, wamount*	*p*-value = 0.1105	*p*-value = 0.02925
(4)	*mprice*	*p*-value = 0.9082	*p*-value = 0.007315
(5)	*mamount*	*p*-value = 0.01344	*p*-value = 0.03328
(6)	*mprice, mamount*	*p*-value = 0.02232	*p*-value = 0.0537
ADL	*price, amount*	*p*-value = 0.9680	*p*-value = 0.6384

Tables 5 and 6 contain the estimation results of the models. The coefficients of high-frequency price and amount variables are calculated based on the estimated weighting parameters in MIDAS models. Consistent with the former results, multiple ADL-MIDAS models except that using monthly transaction data for 1/3 L test set, yield more significant lag coefficients for the high-frequency predictors, thus outperform the single-predictor ADL-MIDAS and the benchmark models. Moreover, evaluated by the AIC statistics, multiple ADL-MIDAS models also have smaller information loss than single-predictor ADL-MIDAS, except MIDAS with *mprice* and *mamount* for 1/3 L forecasting sample. The Jarque-Beta normality test and the Ljung-Box Q test of serial autocorrelation are conducted on residuals. The order of the Ljung-Box Q test is selected as min{floor(n/2)-2,40}. Test results indicate that, for all models and both forecasting samples, residuals satisfy the normality assumption, and are not auto-correlated, on the 5% significance level.

As for the small differences in the prediction performance and model fitting between the weekly predictors and the monthly predictors, it may be closely related to the characteristics of the carbon emissions market in China. That is, ADL-MIDAS models with monthly trading data have slightly smaller *rRMSE* and more significant advantage in prediction accuracy, though less significant predictors. Figure 3 shows the trend of quarterly, monthly, and weekly trading amount data. It can be found that there is usually more transaction at the mid and the end of the year. Generally speaking, the companies are more likely to make the trading decision at a relatively fixed month instead of a certain week. That is, when the companies decide to buy or sell carbon allowances based on their production status in the last period, it might be no difference in trading this week or next week. Another possibility of ineffective information in weekly or even higher frequency data is that the market is not mature yet, and thus the quality of trading data is poor due to the low liquidity, etc. On some trading days, the trading amount is extremely small. And when companies decide to buy or sell carbon allowances, they might fail to find a counterparty due to the low liquidity. That is why the daily trading data are not used as predictors, and weekly trading data may also contain some invalid information.

**Table 5.** Estimation results for GDP forecasting models (1/5 L as test set).

Variables	Weekly Trading Data			Monthly Trading Data			ADL	AR
	(1)	(2)	(3)	(4)	(5)	(6)		
GDP_{t-1}	0.5463 ***	0.914 ***	0.8244 ***	0.5291 ***	0.9444 **	0.7161 **	0.5692 ***	0.7795 *
	(0.1522)	(0.2552)	(0.1042)	(0.1332)	(0.4047)	(0.2613)	(0.0980)	(0.4403)
amount1		−0.0466 *	−0.1255 ***		−0.0799	−0.1951 ***	−0.0281 **	
		(0.0257)	(0.0285)		(0.0483)	(0.0196)	(0.0100)	
amount2		0.3401	0.0657 *		−0.0904	0.1783 ***		
		(0.3250)	(0.0330)		(0.3797)	(0.0189)		
amount3						−0.0365 ***		
						(0.004)		
price1	−0.3272 ***		−0.3569 ***	−0.3283 ***		−0.3846 *	−0.1782 **	
	(0.0494)		(0.0576)	(0.0337)		(0.1718)	(0.0725)	
price2	0.1077 ***		0.0694 **	0.3949 ***		0.2568 *	−0.0983	
	(0.0321)		(0.0257)	(0.1108)		(0.1172)	(0.1025)	
price3						−0.0375 *		
						(0.0153)		
Intercept	3.0393 **	0.5725	1.1507	3.148 ***	0.3695	1.9048	2.9002 ***	1.4795
	(1.0701)	(1.7935)	(0.7278)	(0.9381)	(2.8272)	(1.8125)	(0.688)	(3.1343)
AIC	−30.070	−24.435	−36.400	−25.370	−20.725	−33.177	−25.796	−23.900
JB	0.425	0.665	0.171	0.251	0.705	0.768	0.341	0.724
LB Q	0.612	0.435	0.262	0.481	0.212	0.057	0.355	0.563

Note: The symbols ***, **, and * represent significance under significant levels of 1%, 5%, and 10%, respectively. The standard error of the parameter is in parentheses. JB is the *p*-value for the Jarque-Bera normality test on the residuals, LB Q is the *p*-value for the Ljung-Box Q test of the serial correlation on the residuals.

Table 6. Estimation results for GDP forecasting models (1/3 L as test set).

Variables	Weekly Trading Data			Monthly Trading Data			ADL	AR
	(1)	(2)	(3)	(4)	(5)	(6)		
GDP_{t-1}	0.4229 ***	0.3566 **	0.965 ***	0.4831 *	0.8333 ***	0.8748 ***	0.4354 **	0.5393 ***
	(0.0845)	(0.1534)	(0.0395)	(0.2435)	(0.1052)	(0.1522)	(0.1838)	(0.0786)
amount1		−0.0088 *	0.0864 ***		−0.0661 ***	−0.0683 **	−0.0150	
		(0.0040)	(0.0097)		(0.0155)	(0.0217)	(0.02)	
amount2		-2.62×10^{-5}	−0.0709 ***		5.0838	2.5948		
		(7.06×10^{-4})	(0.0072)		(38.1961)	(5.3568)		
amount3		3.10×10^{-5}	0.0094 ***					
		(2.24×10^{-5})	(0.0009)					
price1	−0.2114 ***		−0.1762 ***	−0.1635 ***		−0.0583	−0.0420	
	(0.0284)		(0.0268)	(0.0399)		(0.0446)	(0.0528)	
price2	0.1148 ***		0.0353 ***	2.6137		−1.9371	−0.1190	
	(0.0200)		(0.0054)	(5.3759)		(3.6016)	(0.0664)	
price3	−0.012 ***		−0.0014 ***					
	(0.0023)		(0.0002)					
Intercept	3.9885 ***	4.4276 ***	0.2296	3.5127 *	1.1876	0.8938	3.8608 **	3.1764 ***
	(0.5974)	(1.0527)	(0.2691)	(1.70000)	(0.7262)	(1.0490)	(1.2845)	(0.5586)
AIC	−32.145	−33.747	−47.471	−28.673	−32.563	−30.175	−26.731	−31.923
JB	0.071	1.363	0.191	0.125	0.805	0.319	0.861	0.651
LB Q	0.883	0.582	0.091	0.392	0.756	0.649	0.642	0.231

Note: The symbols ***, **, and * represent significance under significant levels of 1%, 5%, and 10%, respectively. The standard error of the parameter is in parentheses. JB is the *p*-value for the Jarque-Bera normality test on the residuals, LB Q is the *p*-value for the Ljung-Box Q test of the serial correlation on the residuals.
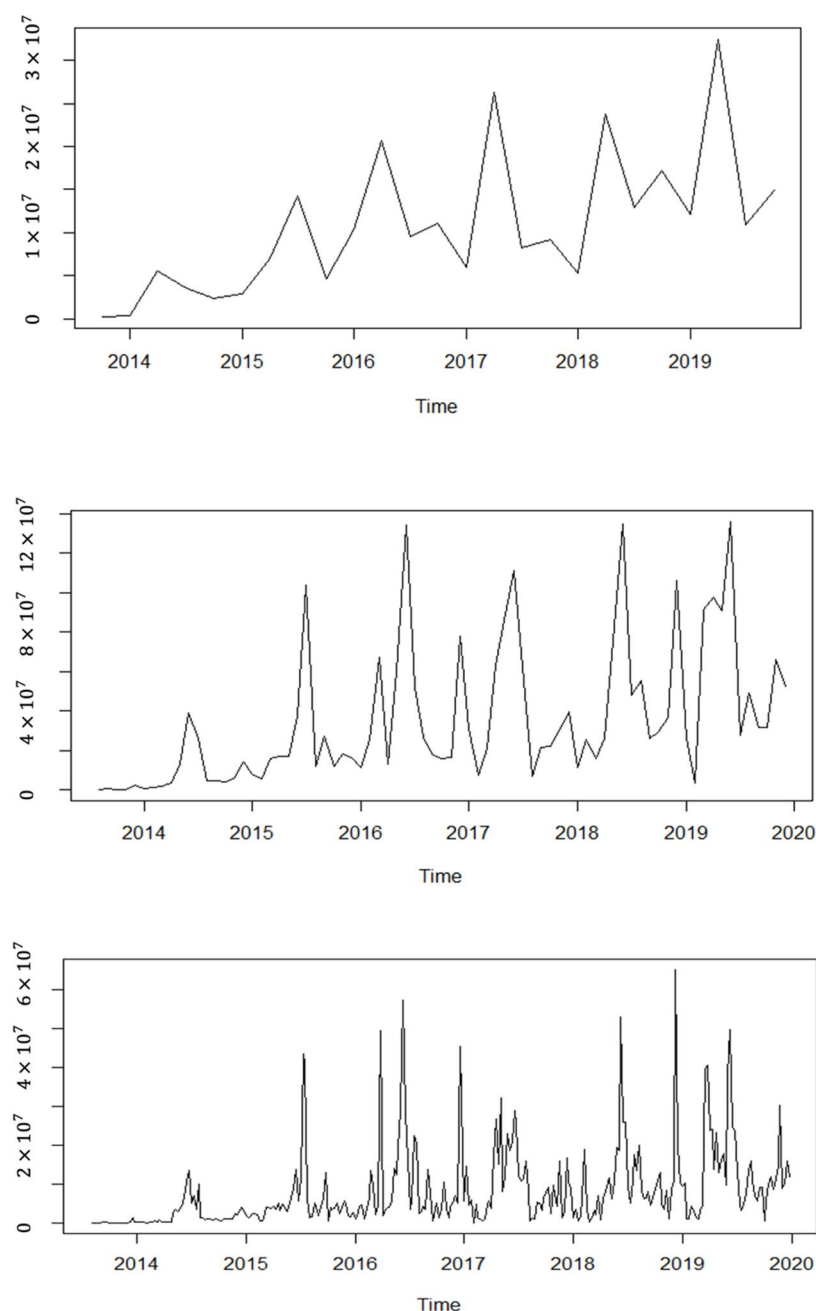
Figure 3. The trend of quarterly, monthly, and weekly trading volume data.

In summary, compared with the benchmark model AR and ADL model using the same frequency trading data, the single-predictor and multiple ADL-MIDAS models using high-frequency carbon emission trading data improve in-sample and out-of-sample nowcasting performance. This indicates that adding trading data might improve quarterly forecasts of real GDP growth, and the MIDAS models could be helpful to make better use of the effective information in high-frequency indicators. As for single-predictor and multiple ADL-MIDAS models, the multiple ADL-MIDAS models using both trading price and trading amount as predictors outperform the single-predictor model with only one predictor. Besides, considering the cycle of carbon trading decision-making and the low liquidity of the current carbon emissions trading data, monthly trading data might be more stable as predictors for quarterly GDP growth forecast than weekly trading data.

Furthermore, we extend the discussion to *h*-step ahead GDP forecast with the multiple ADL-MIDAS. Recall that when *h* = *0*, our model specification becomes a typical nowcast or

real-time forecast; and when predicting the GDP growth of the first quarter, *h = 1/3* step means that the lasted data of predictors starts from February, whereas *h = 2/3* step means that the lasted data of predictors starts from January.

Table 7 shows the nowcast and h-step ahead forecast of multiple ADL-MIDAS models compared to benchmark model AR(1). It is found that when *h < 1*, ADL-MIDAS models have a dominant advantage in the out-of-sample forecast with monthly trading data. That indicates that the MIDAS models perform better in nowcasting. The possible reason is that the high-frequency data has a large amount of real-time transaction information, most of which may become outdated in forwarding prediction.

Table 7. Nowcast and H-step ahead GDP forecast from M-MIDAS (1/3 L as test set).

Model	Forecast Horizon	In-Sample		Out-of-Sample	
		RMSE	*rRMSE*	*RMSE*	*rRMSE*
M-MIDAS (Monthly trading data)	h = 0	0.0384	0.6148	0.2443	0.6582
	h = 1/3	0.0460	0.7361	0.1831	0.4933
	h = 2/3	0.0460	0.7366	0.1796	0.4839
	h = 1	0.0487	0.7806	0.2535	0.6829
	h = 2	0.0599	0.9587	0.7439	2.0043
M-MIDAS (Weekly trading data)	h = 0	0.0367	0.5873	0.2774	0.7473
	h = 1/3	0.0193	0.3095	0.3029	0.8161
	h = 2/3	0.0216	0.3465	0.4622	1.2452
	h = 1	0.0240	0.3842	0.7486	2.0169
	h = 2	0.0292	0.4671	0.4028	1.0852
AR	h = 1	0.0625	1.0000	0.3712	1.0000
	h = 2	0.0778	1.2454	0.5527	1.4890

Last but not least, the combined forecast is examined. We use a combination forecast to explore whether the groups adding high-frequency trading data have better forecasting performance than the groups with only common monthly macroeconomic factors. Besides the usefulness of the combined forecast discussed in Section 3.4, another important reason for choosing a combination forecast is to address the problem of lack of degrees of freedom. For example, adding three factors directly in the multiple ADL-MIDAS models would increase the number of parameters by 6 to 9. And too many parameters will affect the accuracy of the estimation considering the limited sample size.

We use the growth rate of the following three common macroeconomic factors: investment in fixed assets (*cv1*), total retail sales of consumer goods (*cv2*), and the total value of imports and exports (*cv3*). Each of these three control variables is added to the multiple ADL-MIDAS in Equation (5) respectively, as additional predictors besides carbon emission trading variables. As a result, three new multiple ADL-MIDAS models are obtained. Their forecast is subsequently combined weighting by EW, BIC, MSFE, and DMSFE, respectively. Finally, the prediction accuracy of the combined forecast is evaluated and compared to that of single-predictor ADL-MIDAS as well as the benchmark AR model. Similarly, the total dataset was split into a training set and a test set, and 1/5 L and 1/3 L test set lengths are considered respectively. Nowcasting is still the focus, that is, we use the current trading data to predict the GDP growth in current period. However, the latest data of three factors start from one month ago due to the delay in publication time. For example, when predicting the first-quarter GDP growth at the end of March, the latest trading data available starts in March, but the latest data of the chosen macroeconomic factors start in February.

Group (1) is the benchmark group, containing three single-predictor models, each with one monthly macroeconomic factor. Group (2) adds the multiple ADL-MIDAS model with monthly trading data as the fourth model. And Group (3) adds the multiple ADL-MIDAS model with weekly trading data as the fourth model. The weighted average of the predictions from the models within one group is the final forecast from each group.

Tables 8 and 9 show the out-of-sample performance of different groups. It is obvious that all the groups outperform the AR model. As for the four weighting functions in the combination forecast, it can be found that for the 1/5 L test set, both BIC and EFMSE are more efficient weighting function, and for the 1/3 L test set, BIC weighting is the most efficient, by generally generating smallest error measurements. Specifically, for the 1/5 L test set, both Group (2) and Group (3) show smaller *rRMSE* than the benchmark Group (1), especially with BIC weighting. For the 1/3 L test set, the BIC weighting out-of-sample performance of Group (2) and Group (3) stays similar to the benchmark Group (1). But for the other three weighting functions, Group (2) and Group (3) still outperform the benchmark Group (1), indicating that the groups containing high-frequency predictors have more stable forecasting performance. Thus, even taking common and useful macroeconomic predictors into consideration, it is still helpful to introduce high-frequency carbon emission trading data as new predictors.

Table 8. Out-of-Sample performance of combination forecasts for GDP by different groups (1/5 L as test set).

Selection Criteria	*rRMSE*			*rMAPE*			*rMASE*		
GROUP	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
EW	0.993	0.901	0.933	0.876	0.823	0.835	0.866	0.815	0.827
BICW	0.998	0.788	0.799	0.976	0.769	0.715	0.972	0.768	0.712
MSFE	0.924	0.828	0.865	0.834	0.771	0.792	0.825	0.765	0.785
DMSFE	0.923	0.824	0.863	0.833	0.768	0.791	0.823	0.762	0.784
AR	1			1			1		

Table 9. Out-of-Sample performance of combination forecasts for GDP by different groups (1/3 L as test set).

Selection Criteria	*rRMSE*			*rMAPE*			*rMASE*		
GROUP	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
EW	0.882	0.807	0.844	0.868	0.805	0.829	0.867	0.807	0.828
BICW	0.637	0.635	0.747	0.599	0.599	0.711	0.598	0.598	0.713
MSFE	0.784	0.716	0.768	0.768	0.716	0.752	0.767	0.720	0.752
DMSFE	0.786	0.712	0.769	0.770	0.712	0.753	0.769	0.716	0.752
AR	1			1			1		

5.2. Forecasting Results for Monthly Manufacturing PMI Growth

To address possible concerns about the previous results due to the limited sample size, in this section, the monthly PMI growth is used as a new target variable and therefore there are three times more data points. Based on the autocorrelation and partial autocorrelation of PMI growth in Figure 4, we select the AR (1) model as the benchmark model. In practice, we also test AR models with higher lag orders larger and AR (1) is still turned out an appropriate choice considering the significance of the coefficient and the model fitting.

Due to limited space, we only consider 1/3 of the total sample size as the test set. Monthly PMI growth is the target variable, monthly trading data (*mprice*, *mamount*) become the same frequency predictors in the ADL model. Thus, weekly trading data (*wprice*, *wamount*) are used as high-frequency predictors in the single-predictor and multiple ADL-MIDAS models.

Table 10 provides *rRMSE* results for PMI growth forecasts. Compared with benchmark model AR, the multiple ADL-MIDAS model using both high-frequency predictors of the weekly trading price change and trading amount change and the single-predictor ADL-MIDAS model using only weekly trading price change, improve in-sample and out-of-sample prediction performance. The ADL model with the same frequency trading data and the single-predictor ADL-MIDAS model using only weekly trading amount change,

perform well in-sample but fail to outperform the AR model out-of-sample. Besides, the multiple ADL-MIDAS model consistently yields smaller *rRMSE* than the single-predictor ADL-MIDAS model. This conclusion is similar to that of the quarterly GDP growth forecast.
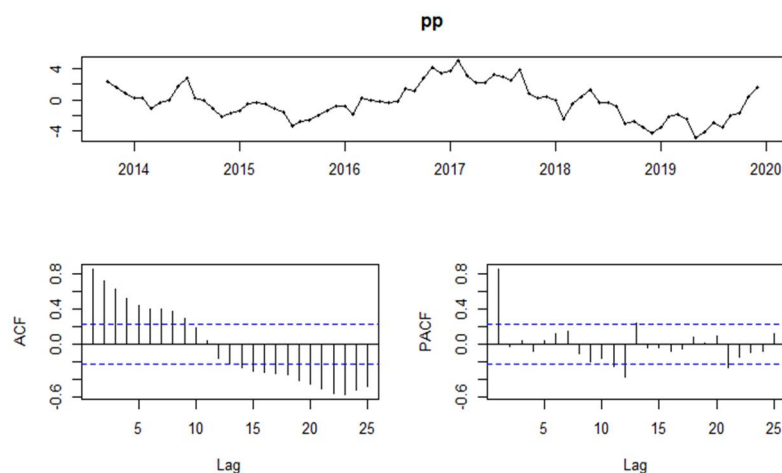

Figure 4. Autocorrelation and partial autocorrelation graph of PMI growth.

Table 10. PMI nowcasting performance comparison by *rRMSE*.

Model	Independent Variables	In-Sample	Out-of-Sample		
			Fixed Window	Rolling Window	Recursive Window
MIDAS	*wprice*	0.9456	0.9657	0.9860	0.9803
MIDAS	*wamount*	0.8593	1.0811	1.0481	1.0700
M-MIDAS	*wprice, wamount*	0.7733	0.8916	0.9255	0.8973
ADL	*mprice, mamount*	0.8480	1.0561	1.0702	1.0587
AR		1	1	1	1

Table 11 shows the Diebold-Mariano test results. Although the out-of-sample *rRMSE* of the ADL model using the same frequency trading data is slightly bigger than 1, the *p*-value of the Diebold-Mariano test statistic is 0.64 and we can't reject the null hypothesis that the two benchmarks, the ADL model and the AR model, have equal prediction accuracy. For the multiple ADL-MIDAS model using weekly trading data, the Diebold-Mariano test results show that it is significantly better than the ADL model (*p*-value of 0.06 rejects the null hypothesis of equal prediction accuracy at the 10% significance level), but not significantly better than the AR model (*p*-value of 0.18 fails to reject the null hypothesis of equal prediction accuracy at the 10% significance level).

Table 11. PMI nowcasting performance comparison by Diebold-Mariano test (fixed window).

Test Model	Predictors	H$_1$: Test Model Is More Accurate than Benchmark Model	
		AR as Benchmark Model	ADL as Benchmark Model
MIDAS	*wprice*	*p*-value = 0.3534	*p*-value = 0.2215
MIDAS	*wamount*	*p*-value = 0.7532	*p*-value = 0.5853
M-MIDAS	*wprice, wamount*	*p*-value = 0.1765	*p*-value = 0.06173

It is not surprising that monthly trading data might be a more stable predictor for quarterly GDP growth forecast than weekly trading data, while weekly trading data is a better predictor for monthly PMI growth. The possible reason might be that the information contained in the monthly carbon emission trading data often correspond to the production status for a longer period. On the other hand, the companies are more likely to make the trading decision at relatively fixed months at the mid and the end of the year based on

their production status in the last period. So, weekly trading data might contain more information related to the current month, while monthly data contains information for a longer period.

　　Table 12 provides the results of coefficient estimation. Compared with the benchmark models AR, ADL, and the single-predictor MIDAS models, the multiple ADL-MIDAS model has more significant coefficients. This is consistent with the out-of-sample forecasting performance. However, evaluated by AIC, the ADL-MIDAS with *wamount* turns out to have the least information loss among the five models. Jarque-Bera normality test results again confirm that the residuals from all models satisfy the normality assumption. Ljung-Box Q test results show that, except for the ADL-MIDAS with *wamount*, all the other models have no auto-correlated residuals.

Table 12. Estimation results for PMI forecasting models.

Variables	(1)	(2)	(3)	ADL	AR
PMI_{t-1}	0.8987 *** (0.0689)	0.8941 *** (0.8941)	0.8811 *** (0.0675)	0.8758 *** (0.0779)	0.8850 *** (0.0709)
price1	−1.4180 ** (0.6070)		−0.1204 (0.0954)	−0.7711 (0.7990)	
price2	0.4436 (0.4345)		0.0631 ** (0.0249)	−1.0234 (0.6381)	
price3			−0.0049 *** (0.0016)		
amount1		−0.1241 (0.0899)	2.7905 * (1.4373)	−0.0969 (0.1870)	
amount2		0.0615 ** (0.0244)	−2.2288 * (1.0968)	0.3283 ** (0.1455)	
amount3		−0.0048 *** (0.0015)	0.3527 ** (0.1693)	−0.3464 ** (0.1304)	
Intercept	−0.2650 (0.1897)	0.2029 (0.1670)	0.1287 (0.1252)	−0.2800 (0.2088)	0.0896 (0.1447)
AIC	104.726	97.121	101.711	106.353	113.900
JB	0.260	0.184	0.627	0.413	0.658
LB Q	0.294	0.004	0.201	0.185	0.257

Note: The symbols ***, **, and * represent significance under significant levels of 1%, 5%, and 10%, respectively. The standard error of the parameter is in parentheses. JB is the *p*-value for the Jarque-Bera normality test on the residuals, LB Q is the *p*-value for the Ljung-Box Q test of the serial correlation on the residuals.

　　Besides, we select different estimation windows in this part. It seems that the fixed window has better out-of-sample prediction performance than the recursive window or rolling window, though the difference is not obvious. This might be related to the algorithm settings. For recursive window and rolling window, the parameters would be recalculated with new data, but the lag orders and the weight schemes are determined at the beginning.

　　Since there more observations for PMI than GDP in this sample period, the U-MIDAS model without constraints is also taken into consideration. Table 13 shows the in-sample and out-of-sample *rRMSE* of the multiple ADL-MIDAS models and the corresponding U-MIDAS models with the same lag but no weighting scheme. As mentioned before, we select three types of weighting schemes: Almon polynomial (almonp), exponential Almon polynomial (nealmon), and Beta distribution polynomial (nbeta). For the two high-frequency predictors in multiple models, we get nine candidate ADL-MIDAS models with different combinations of weighting schemes. It can be found that all nine candidate ADL-MIDAS models have smaller out-of-sample *RMSE* than the benchmark model AR, indicating that adding high-frequency trading data is indeed helpful. Specifically, the multiple ADL-MIDAS model in the previous analysis is the fifth ADL-MIDAS model in Table 13, and it is selected by AIC, in-sample *RMSE*, and significance of the coefficients. The U-MIDAS models show advantages in-sample, but the ADL-MIDAS models with weighting schemes have smaller out-of-sample *RMSE*. It is consistent with the discussion

of the weighting scheme in Section 3. Although the U-MIDAS model can retain the information in the data to the largest extent, excessive consumption of degrees of freedom might cause considerable efficiency loss. What's more, there are nine different combinations of weighting schemes and it could be the case that none of them constitute an appropriate constraint, so even incorrect constraints might be useful when the sample size is small.

Table 13. PMI nowcasting performance of M-MIDAS and U-MIDAS by *rRMSE*.

	In-Sample		Out-of-Sample		Weighting Scheme
	M-MIDAS	**U-M-MIDAS**	**M-MIDAS**	**U-M-MIDAS**	
1	0.8829	0.7514	0.8531	1.0743	nealmon & nealmon
2	0.8272	0.4164	0.9773	2.3280	nealmon & almonp
3	0.8829	0.8267	0.8530	0.8819	nealmon & nbeta
4	0.8048	0.6797	0.9853	1.0934	almonp & nealmon
5	0.7733	0.6307	0.8916	1.4749	almonp & almonp
6	0.9082	0.7481	0.9881	1.1403	almonp & nbeta
7	0.8712	0.7234	0.8607	1.2745	nbeta & nealmon
8	0.8426	0.6797	0.9658	1.0934	nbeta & almonp
9	0.9083	0.5327	0.9004	1.1936	nbeta & nbeta

In sum, the results of using trading data to forecast monthly manufacturing PMI growth show that the multiple ADL-MIDAS models with high-frequency carbon emission trading data improve the out-of-sample prediction accuracy beyond the AR model and ADL model using the same frequency trading data. And the conclusion is consistent with the previous results in the forecast of quarterly GDP growth.

6. Conclusions

Since the establishment of the pilot carbon emissions trading in 2013, the high-frequency carbon emissions trading data has been available. This paper provides a preliminary analysis and discussion on the quality of carbon emissions trading data as a predictor for macroeconomic indicators. With the national carbon emissions trading system of China officially launched in 2021, the quality of carbon trading data will continue to improve. Existing research of China's carbon market focuses on the pricing mechanism, thus our work fills the gap in the area of the application of carbon trading data. Containing information about the state of the economy and available on a daily basis without time lag, the high-frequency carbon emission trading data have the potential to be good predictors in macro forecasts.

The method chosen in this study is MIDAS, which can directly deal with data sampled at different frequencies and retaining the effective information in the high-frequency indicators. Monthly and weekly trading price change and trading amount change of carbon emissions market are chosen as high-frequency indicators to establish single-predictor and multiple ADL-MIDAS models, to predict quarterly GDP growth rate and monthly PMI growth rate. The forecasting performance of competing models is evaluated with both *rRMSE* and Diebold-Mariano test. By comparing with a benchmark AR model, we find models with high-frequency carbon emission trading data have better forecasting performance than models without trading data. By comparing with the ADL model using the same frequency carbon emission trading data, we find models with high-frequency carbon emission trading data have better forecasting performance than models with the same frequency trading data. Combination forecasting methods are also adopted to predict GDP growth rate, the results find the combination forecasts from the groups containing high-frequency trading data have better forecasting performance than the groups with only common macroeconomic factors. Thus, it is still useful to introduce high-frequency carbon emission trading data as new predictors; even same-frequency macroeconomic predictors have already done a good job. Further analysis of the nowcast and h-step ahead forecast with MIDAS models, we found that the advantage of MIDAS models lies in nowcasting.

The possible reason is that the high-frequency data has a large amount of information and therefore has more outdated information in forwarding prediction.

Concerning the forecast of monthly manufacturing PMI growth rate, it can be shown that the multiple ADL-MIDAS models with high-frequency weekly carbon emission trading data improve the out-of-sample prediction accuracy beyond the AR model and the ADL model using the same frequency trading data. The conclusion is consistent with the previous results in the forecast of quarterly GDP growth.

The results of this study can be improved with more data availability and higher data quality after the official launch in the following ways. Firstly, there will be more flexible choices of multiple models and higher credibility of statistical results. Secondly, with enhanced liquidity in a more actively traded carbon market, daily trading data may be employed as a predictor. Thirdly, the analysis of this study is based on public information, with more companies trading in the market; it is possible to visit companies to collect detailed information of carbon trading decisions. Last but not least, the macro data are obtained from the official website of the National Bureau of Statistics of China. Considering the problem of data revision, using real-time data might be more practical for nowcasting.

References

1. Giannone, D.; Reichlin, L.; Small, D. Nowcasting: The real-time informational content of macroeconomic data. *J. Monet. Econ.* **2008**, *55*, 665–676. [CrossRef]
2. Estrella, A.; Mishkin, F.S. Predicting U.S. recessions: Financial variables as leading indicators. *Rev. Econ. Stat.* **1998**, *80*, 45–61. [CrossRef]
3. Zhang, Y.; Zhang, S. The impacts of GDP, trade structure, exchange rate and FDI inflows on China's carbon emissions. *Energy Policy* **2018**, *120*, 347–353. [CrossRef]
4. Li, H.; An, H. How does the coal stock market, carbon market and coal price co-movement with each other in China: A co-movement matrix transmission network perspective. *Energy Procedia* **2017**, *105*, 3479–3484. [CrossRef]
5. Zeng, S.; Nan, X.; Liu, C.; Chen, J. The response of the Beijing carbon emissions allowance price (BJC) to macroeconomic and energy price indices. *Energy Policy* **2017**, *106*, 111–121. [CrossRef]
6. Alberola, E.; Chevallier, J.; Chèze, B. Price drivers and structural breaks in European carbon prices 2005–2007. *Energy Policy* **2008**, *36*, 787–797. [CrossRef]
7. Chevallier, J. Carbon futures and macroeconomic risk factors: A view from the EU ETS. *Energy Econ.* **2009**, *31*, 614–625. [CrossRef]
8. Zou, Y.; Wei, W. The study on the impact factors of certified carbon emissions (CERs) spot markets (in Chinese). *J. Financ. Res.* **2013**, *10*, 142–153.
9. Zhou, T.; Xu, R. The formation and volatility features of China's carbon emission trading price—Based on the data of Shenzhen carbon emission exchange (in Chinese). *J. Financ. Dev. Res.* **2016**, *1*, 16–25.
10. Ghysels, E.; Santa-Clara, P.; Valkanov, R. The MIDAS touch: Mixed data sampling regression models. *Cirano Work. Pap.* **2004**, *5*, 512–517. Available online: http://www.cirano.qc.ca/files/publications/2004s-20.pdf (accessed on 6 February 2021).
11. Ghysels, E.; Santa-Clara, P.; Valkanov, R. There is a risk-return trade-off after all. *J. Financ. Econ.* **2005**, *76*, 509–548. [CrossRef]
12. Ghysels, E.; Santa-Clara, P.; Valkanov, R. Predicting volatility: Getting the most out of return data sampled at different frequencies. *J. Econom.* **2006**, *131*, 59–95. [CrossRef]
13. Ghysels, E.; Valkanov, R. Linear Time Series Processes with Mixed Data Sampling and MIDAS Regression Models. Available online: http://ssrn.com/abstract=920610 (accessed on 6 February 2021).

14. Ghysels, E.; Sinko, A.; Valkanov, R. MIDAS regressions: Further results and new directions. *Econom. Rev.* **2007**, *26*, 53–90. [CrossRef]

15. Andreou, E.; Ghysels, E.; Kourtellos, A. Regression models with mixed sampling frequencies. *J. Econ.* **2010**, *158*, 246–261. [CrossRef]

16. Marcellino, M.; Schumacher, C. *Factor Nowcasting of German GDP with Ragged-Edge Data: A Model. Comparison Using MIDAS Projections*; Deutsche Bundesbank Discussion Paper, Series 1, No. 34; Deutsche Bundesbank Publishing: Frankfurt, Germany, 2007; Available online: https://www.bankofcanada.ca/wp-content/uploads/2010/09/schumacher.pdf (accessed on 6 February 2021).

17. Breitung, J.; Roling, C. Forecasting inflation rates using daily data: A nonparametric MIDAS Approach. *J. Forecast.* **2015**, *34*, 588–603. [CrossRef]

18. Foroni, C.; Marcellino, M.; Schumacher, C. Unrestricted mixed data sampling (MIDAS): MIDAS regressions with unrestricted lag polynomials. *J. R. Stat. Soc. Ser. A (Stat. Soc.)* **2015**, *178*, 57–82. [CrossRef]

19. Barsoum, F.; Stankiewicz, S. Forecasting GDP growth using mixed-frequency models with switching regimes. *Int. J. Forecast.* **2015**, *31*, 33–50. [CrossRef]

20. Clements, M.P.; Galvão, A.B. Macroeconomic forecasting with mixed frequency data: Forecasting US output growth and inflation. *J. Bus. Econ. Stat.* **2008**, *26*, 546–554. [CrossRef]

21. Ghysels, E.; Wright, J.H. Forecasting professional forecasters. *J. Bus. Econ. Stat.* **2009**, *27*, 504–516. [CrossRef]

22. Clements, M.P.; Galvão, A.B. Forecasting US output growth using leading indicators: An appraisal using MIDAS models. *J. Appl. Econ.* **2009**, *24*, 1187–1206. [CrossRef]

23. Kuzin, V.; Marcellino, M.; Schumacher, C. MIDAS vs. mixed-frequency VAR: Nowcasting GDP in the euro area. *Int. J. Forecast.* **2011**, *27*, 529–542. [CrossRef]

24. Bai, J.; Ghysels, E.; Wright, J.H. State space models and MIDAS regressions. *Econ. Rev.* **2013**, *32*, 779–813. [CrossRef]

25. Andreou, E.; Ghysels, E.; Kourtellos, A. Should macroeconomic forecasters look at daily financial data and how? *J. Bus. Econ. Stat.* **2013**, *31*, 240–251. [CrossRef]

26. Bessec, M.; Bouabdallah, O. Forecasting GDP over the business cycle in a multi-frequency and data-rich environment. *Oxf. Bull. Econ. Stat.* **2015**, *77*, 360–384. [CrossRef]

27. Ghysels, E.; Kvedaras, V.; Zemlys, V. Mixed frequency data sampling regression models: The R Package MIDASR. *J. Stat. Softw.* **2016**, *72*, 1–35. [CrossRef]

28. Sinko, A. Some Applications of Mixed Data Sampling Regression Models. Ph.D. Dissertation, Department of Economics, the University of North Carolina at Chapel Hill, Chapel Hill, NC, USA, 2008. Available online: https://cdr.lib.unc.edu/concern/file_sets/bg257f55p (accessed on 6 February 2021).

29. Stock, J.H.; Watson, M.W. Combination forecasts of output growth in a seven-country data set. *J. Forecast.* **2004**, *23*, 405–430. [CrossRef]

30. Hendry, D.F.; Clements, M.P. Pooling of forecasts. *Econ. J.* **2004**, *7*, 1–31. [CrossRef]

31. Timmermann, A. *Forecast. Combinations. Handbook of Economic Forecasting*; Elliott, G., Granger, C., Timmermann, A., Eds.; Elsevier: Amsterdam, The Netherlands, 2006; Volume 1, pp. 136–196. [CrossRef]

32. Aiolfi, M.; Timmermann, A. Persistence in forecasting performance and conditional combination strategies. *J. Econ.* **2006**, *135*, 31–53. [CrossRef]

33. Andreou, E. On the use of high frequency measures of volatility in MIDAS regressions. *J. Econ.* **2016**, *193*, 367–389. [CrossRef]

34. Diebold, F.X.; Mariano, R.S. Comparing predictive accuracy. *J. Bus. Econ. Stat.* **1995**, *13*, 253–263. [CrossRef]