FORECASTING OUTPUT GROWTH AND INFLATION WITH STOCK RETURNS: A MIDAS APPROACH

YU   CAI

SINGAPORE MANAGEMENT UNIVERSITY

2012

Forecasting Output Growth and Inflation with Stock Returns:
A MIDAS Approach

by

Yu   Cai

Submitted to School of Economics in partial fulfillment

of the requirements for the Degree of Master of Science in Economics

Thesis Committee:

Anthony S. Tay (Supervisor/Chair)

Associate Professor of Economics

Singapore Management University

Yu Jun

Professor of Economics and Finance

Singapore Management University

Tse Yiu Kuen

Professor of Economics

Singapore Management University

Singapore Management University

2012

Forecasting Output Growth and Inflation with Stock Returns:

A MIDAS Approach

YU CAI

SCHOOL OF ECONOMICS

SINGAPORE MANAGEMENT UNIVERSITY

90 STAMFORD ROAD

SINGAPORE 178903

TEL: (65) 8335_7959

2012

Abstract

Improving forecasts of macroeconomic indicators such as output growth and inflation is of focal interest to academics and policy makers. Because stock returns can be observed at very high frequencies, there is the question of whether high frequency information is useful for forecasting output and inflation. Furthermore, stock data is timely, whereas macro data is available only at a lag, so there is a question of whether stock returns can help to indicate the current state of the economy, i.e., "nowcasting". In this thesis, we study the predictive power of daily stock returns on output growth and inflation with **Mi**xed **Da**ta **S**ampling (henceforth, MIDAS) regression models both in forecasting and nowcasting contexts. We filter the daily stock returns with a newly proposed frequency domain filter, and aggregate the daily data with MIDAS weights using estimated parameter values. We find that predictors with MIDAS regressions perform quite well in inflation forecasting. For Singapore inflation, filtered stock returns forecast better than unfiltered stock returns; for US inflation, on the other hand, unfiltered stock returns forecast better than filtered stock returns. Predictors with MIDAS regressions perform fairly well in Singapore output growth forecasting in that contemporary stock returns have higher forecasting accuracy than the benchmark model, but for the US output growth, we don't see any improvements with our MIDAS regressions.

Keywords: Forecasting; Nowcasting; MIDAS; Frequency Domain Filter; Diebold-Mariano Test

Table of Contents

1 Introduction

Forecasting macroeconomic variables is an important task for central banks, financial firms, and any other entities whose outcome depend on business cycle conditions. Unfortunately, many important macroeconomic indicators are not sampled at the same frequency. The availability of data sampled at different frequencies always presents a dilemma for researchers. On the one hand, the variables that are available at high-frequencies contain potentially valuable information. On the other hand, the standard time series models cannot use this high-frequency information directly if some of the variables are available at a lower frequency, because they are designed for use with data sampled at the same interval. The common solution in such cases is to aggregate the high frequency data so that the left-hand and right-hand variables are available at the same frequency. In the process, a lot of potentially useful information might be discarded, thus rendering the relation between the variables difficult to detect. As an alternative, Ghysels et.al.(2004, 2005, 2006), have proposed regressions that directly accommodate variables sampled at different frequencies. Their Mixed Data Sampling (MIDAS) regressions represent a simple, parsimonious, and flexible class of time series models that allow the left-hand and right-hand variables of time series regressions to be sampled at different frequencies. Although high-frequency data contains potentially useful information, it may also contain noises as well as high frequency components that may not be useful for forecasting the lower frequency macroeconomic variables. Therefore, we propose to filter the high-frequency data first and figure out exactly which cycle is most useful for forecasting the output growth.

Besides, using standard time series regression models where the regressors are aggregated to some low frequency, such as financial aggregates (that are available at higher

frequencies), can also yield estimation problems. Andreou, Ghysels, and Kourtellos (2010) show that the estimated slope coefficient of a regression model that imposes a standard equal weighting aggregation scheme (and ignores the fact that processes are generated from a mixed data environment) yields asymptotically inefficient (at best) and in many cases inconsistent estimates. As is well known, both inefficiencies and inconsistencies can have adverse effects on forecasting.

The gains of real-time forecast updating, sometimes called nowcasting when it applies to current quarter assessments, have also been of particular interest to policy makers. The simplicity of the MIDAS approach allows us to produce nowcasts with potentially a large set of real-time high frequency data feeds. More importantly, the MIDAS regressions can be extended beyond nowcasting the current quarter to produce direct forecasts multiple quarters ahead.

Before we proceed, it is useful to introduce a simple MIDAS regression here. Suppose that a variable y_t is available once between $t-1$ and t (say, quarterly), another variable $x_t^{(m)}$ is observed m times in the same period (say, daily or $m = 66$), and that we are interested in the dynamic relation between y_t and $x_t^{(m)}$. In other words, we want to project the left-hand variable y_t onto a history of lagged observations of $x_{t-j/m}^{(m)}$. The superscript on $x_{t-j/m}^{(m)}$ denotes the higher sampling frequency, and its exact timing lag is expressed as a fraction of the unit interval between $t-1$ and t. A simple MIDAS model is:

$$Y_t = \beta_0 + \beta_1 B(L^{1/m}; \theta) x_t^{(m)} + \varepsilon_t^{(m)} \tag{1}$$

for $t = 1, \ldots, T$, where $B(L^{1/m}; \theta) = \sum_{k=0}^{K} B(k; \theta) L^{k/m}$ and $L^{1/m}$ is a lag operator such that $L^{1/m} x_t^{(m)} = x_{t-1/m}^{(m)}$; the lag coefficients in $B(k; \theta)$ of the corresponding lag operator $L^{k/m}$ are parameterized as a function of a small-dimensional vector of parameters θ.

In the mixed-frequency framework (1), the number of lags of $x_t^{(m)}$ is likely to be significant. For instance, if quarterly observations of y_t is affected by six months' worth of lagged $x_t^{(m)}$'s, we would need 132 lags ($K = 132$) of high-frequency lagged variables. If the parameters of the lagged polynomial are left unrestricted (or $B(k)$ does not depend on θ), then there would be a lot of parameters to estimate. As a way of addressing parameter proliferation, in a MIDAS regression the coefficients of the polynomial in $L^{1/m}$ are captured by a known function $B(k; \theta)$ of a few parameters summarized in a vector θ. We will discuss two alternative specifications of $B(k; \theta)$ in this thesis. Finally, the parameter β_1 captures the overall impact of lagged $x_t^{(m)}$'s on y_t .

The rest of the thesis is organized as follows: A literature review is presented in Section 2. The MIDAS methodology is introduced and discussed in detail in Section 3. Section 4 describes the data we apply for forecasting, with particular emphasis on parameter estimation of the MIDAS regressions and the Ouliaris and Corbae (2005)'s frequency domain filter. Section 5 presents our forecasting results for both the unfiltered and filtered daily stock returns, we also introduce the Diebold-Mariano test and its application to compare the predictive accuracy of different predictors. Section 6 concludes.

2 Literature Review

The MIDAS regression models were introduced by Ghysels et.al. (2004). They examine the asymptotic properties of MIDAS regression estimation and compare it with traditional distributed lag models. They show that MIDAS regressions will always lead to more efficient estimation than the typical approach of aggregating all series to the least frequent sampling, and in some cases MIDAS regressions are also as efficient as distributed lag regressions with

all series available at the highest frequency.

Since then, a large literature has focused on applying MIDAS in volatility forecasting. For instance, Ghysels et.al.(2005) uses monthly and daily market returns data from 1928 to 2000 and with MIDAS as a model of the conditional variance, they show that there is a significantly positive relation between market volatility and return (expected returns are proxied using monthly averages, while the variance is estimated using daily squared returns over the last year). They also include business cycle variables together with both the symmetric and asymmetric MIDAS estimators of conditional variance in the ICAPM (Intertemporal Capital Asset Pricing Model) equation and find that the trade-off between risk and return is virtually unchanged, and the explanatory power of the conditional variance for expected returns is not affected by the inclusion of other predictive variables.

Ghysels et.al.(2006) studies the predictability of return volatility with MIDAS regressions, the MIDAS regression framework allows them to investigate whether the use of high-frequency data necessarily leads to better volatility forecasts at various horizons and allows for a great degree of flexibility. They find that daily realized power is the best predictor of future increments in quadratic variation, followed by the daily range. The prediction equations involve about 50 daily lags and find that there is no real benefit of using intra-daily data directly. The MIDAS regressions outperform other linear forecast models involving daily realized volatility. Ghysels et.al.(2007) presents various extensions of MIDAS models, such as a generalized MIDAS regression, nonlinear MIDAS regressions, tick-by-tick MIDAS regressions, and multivariate MIDAS regressions. They apply some of the generalizations to estimate the relation between conditional expected return and risk using ten years of daily Dow Jones index return data.

Alper et.al.(2008) explores the relative weekly stock market volatility forecasting performance of the linear univariate MIDAS regression model based on squared daily returns vis--vis the benchmark model of GARCH(1,1) for a set of four developed and ten emerging market economies. Their findings show that the MIDAS squared daily return regression model outperforms the GARCH model significantly in four of the emerging markets. Moreover, the GARCH model fails to outperform the MIDAS regression model in any of the emerging markets significantly, while the results are slightly less conclusive for the developed economies.

Alper et.al.(2011) evaluates weekly out-of-sample volatility forecast performance of univariate MIDAS model compared to the benchmark model of GARCH(1,1) for ten emerging stock markets. The results show that the MIDAS model offers statistically better forecasting precision during the recent financially turbulent era, based on the test suggested by West(2006). For the tranquil period, however, the MIDAS model can not produce statistically better weekly volatility forecast.

A few other authors also apply the MIDAS approach to forecast output growth by using high frequency data. For instance, Tay (2006) considers augmenting the quarterly AR(1) model for real output growth with daily returns using MIDAS methods and augmenting the quarterly AR(1) model with the most recent r-day returns as an additional predictor, and finds that for the early 2000s, his MIDAS model outperforms his benchmark model by 20 to 30 percent, while his AR model using stock returns over a specified period performs even better. Clements and Galvao (2008) extends the distributed-lag MIDAS specification to include an AR term (MIDAS-AR) and finds that the use of monthly data on the current quarter leads to significant improvement in forecasting current and next

quarter output growth. The use of real-time vintage data serves to strengthen their finding that within-quarter monthly indicator information can result in marked improvements in forecast performance. MIDAS fares well relative to the other models that use monthly information. Coupled with its flexibility and ease of use relative to methods that involve generation of forecasts of explanatory variables offline, the MIDAS-AR would appear to be a useful addition to the sets of models and methods that exploit monthly indicators for the short-term forecasting of macroaggregates. Kuzin, Marcellino, and Schumacher(2011) use monthly series to forecast euro-area quarterly GDP. They compared the performance of the AR-MIDAS model of Clement and Galvao (2008) to a vector autoregression (VAR) and find that the AR-MIDAS model performs better near one-quarter horizons, while the VAR model performs better near three-quarter horizon. Galvao (2011)include a MIDAS framework in a smooth-transition autoregression to allow for changes in a higher-frequency variable's forecasts of quarterly GDP when using weekly short-term interest rate and stock returns data along with term spread data, sometimes up to horizon of two or three years.

Clements and Galvao (2009) considers the MIDAS regression model as a way of combining the information in multiple leading indicators to predict output growth up to 1 year ahead. They find that: (a) MIDAS is a useful vehicle for combining a small group of indicators for forecasting; (b) the use of information on the current quarter improves forecasts and (c) combination in modeling with MIDAS is better than combination of forecast when predicting the direction of change of output growth.

As for the inflation forecasting, Ghysels et.al.(2009) offers some insights about using MIDAS methods. They show that for forecasting CPI inflation the univariate MIDAS model yields forecasting gains for one quarter ahead of about 85%, 53%, and 19% over the

RW, AR and FAR, respectively. Interestingly, for longer forecasting horizons of 8 quarters ahead they find that the best MSFE given by the parsimonious univariate MIDAS model yields 28% forecasting gains over the RW and AR and around 50% gains over the traditional Factor models.

Compared to the previous research, our study makes three new contributions. First, we filter the daily stock returns using a newly proposed frequency domain filter, which enables us to figure out which cycles of the data contains potential useful information for predicting output growth and inflation; Second, our finding shows that the MIDAS methodology is more effective in forecasting inflation than in output growth; Last but not the least, we compare the predictive accuracy of different forecasts by applying Diebold-Mariano test and dig out some new insights about the predictors.

3   MIDAS

The MIDAS approach models the response of the dependent variable to the higher-frequency explanatory variables as a highly parsimonious distributed lag, as a way of preventing the proliferation of parameters that might otherwise result. Modeling the coefficients on the lagged explanatory variables as a distributed lag function allows for long lags with only a small number of parameters needing to be estimated. They allow us to study, in a unified framework, the forecasting performance of a large class of models which involve: (i) data sampled at different frequencies; (ii) various past data window lengths; and (iii) different regressors. The specification of the regressions combine recent developments regarding estimation of volatility and a not so recent literature on distributed lag models.

In this section, we focus on the specification of (1), we deal with finite one-sided poly-

nomials applied to a single regressor. This is one of the simplest MIDAS specifications and it allows us to focus on the parametrization of $B(k;\theta)$. We focus on two parameterizations of $B(k;\theta)$. The first is

$$B(k;\theta) = \frac{exp(\theta_1 k + \cdots + \theta_T k^T)}{\sum_{k=1}^{K} exp(\theta_1 k + \cdots + \theta_T k^T)} \tag{2}$$

which we refer to as the exponential Almon lag. The function $B(k;\theta)$ is known to be quite flexible and can take various shapes with only a few parameters. Ghysels et.al.(2005) use the functional form (2) with two parameters, or $\theta = [\theta_1;\theta_2]$. Figure 1 illustrates the flexibility of the exponential Almon lag even in this simple two-parameter case. First, it is easy to see that for $\theta_1 = \theta_2 = 0$, we have equal weights (this case is not plotted). Second, the exponential function (2) can produce hump shapes as shown in the upper left panel of Figure 1($\theta_1 = 3 * 10^{-2}$ and $\theta_2 = -9 * 10^{-4}$). Third, the weights can decline fast(upper right panel, with $\theta_1 = 7 * 10^{-4}$ and $\theta_2 = -1 * 10^{-3}$) and slowly(bottom left panel, with $\theta_1 = 7 * 10^{-3}$ and $\theta_2 = -1 * 10^{-4}$) with the lag. Finally, it can also produce bell shapes like normal distribution as shown in the bottom right panel($\theta_1 = 3 * 10^{-1}$ and $\theta_2 = -9 * 10^{-4}$)). A declining weight is guaranteed as long as $\theta_2 \leq 0$. It is important to point out that the rate of decline determines how many lags are included in regression (1). Since the parameters are estimated from the data, once the functional form of $B(k;\theta)$ is specified, the lag length selection is purely data driven.

The second parametrization has also only two parameters, or $\theta = [\theta_1;\theta_2]$:

$$B(k;\theta_1,\theta_2) = \frac{f(k/K,\theta_1;\theta_2)}{\sum_{k=1}^{K} f(k/K,\theta_1;\theta_2)} \tag{3}$$

where:

$$f(i,\theta_1,\theta_2) = \frac{i^{\theta_1-1}(1-i)^{\theta_2-1}\Gamma(\theta_1+\theta_2)}{\Gamma(\theta_1)\Gamma(\theta_2)}$$
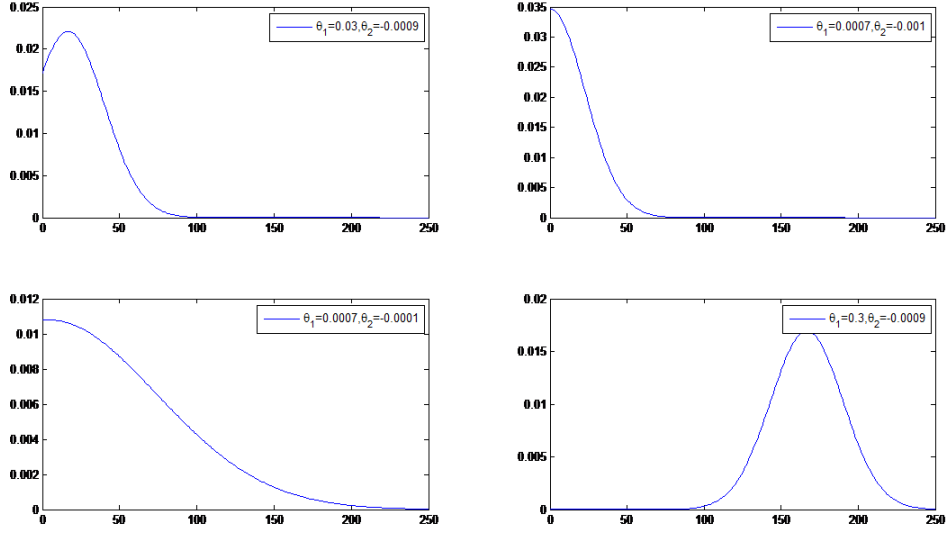
8

Figure 1: Exponential Almon polynomial MIDAS weights. The figure shows various shapes of the exponential Almon specification (2). We plot the weights on the first 252 lags (which correspond to one year's worth of daily lags).
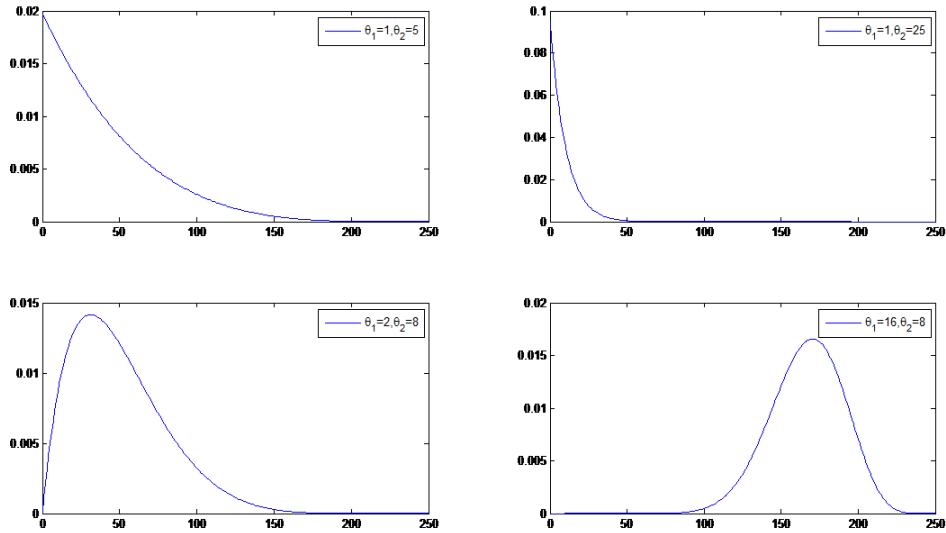

Figure 2: Beta polynomial MIDAS weights. The figure shows various shapes of the beta specification (3). We plot the weights on the first 252 lags (which correspond to one year's worth of daily lags).

$$\Gamma(\theta) \;\; = \;\; \int_0^\infty e^{-x}x^{\theta-1}dx$$

Figure 2 displays various shapes of (3) for several values of θ_1 and θ_2. The function can also take many shapes not displayed in the figure. For instance, it is easy to show that for $\theta_1 = \theta_2 = 1$ we have equal weights (this case is not shown). Second, the upper left panel in Figure 2 shows the case of slowly declining weight that corresponds to $\theta_1 = 1$ and $\theta_2 = 5$. As θ_2 increases, we obtain faster declining weights, as shown in the upper right panel of the figure($\theta_1 = 1$ and $\theta_2 = 25$). Third, the bottom left panel illustrates a hump-shaped pattern which emerges for $\theta_1 = 2$ and $\theta_2 = 8$. Finally, the Beta polynomial function (3) can also produce bell shapes as shown in the bottom right panel($\theta_1 = 16$ and $\theta_2 = 8$). The flexibility of the beta function is well known. As pointed out in the exponential Almon lag case, the rate of weight decline determines how many lags are included in the MIDAS regression.

4 Data

We use a dataset of mixed frequencies (daily and quarterly). Our purpose here is to use daily stock returns to forecast quarterly GDP growth and inflation. For instance, we apply unfiltered and filtered daily returns of Dow Jones Industrial Average(DJI) from January 1, 1951 to December 31, 2010, to forecast the US GDP growth and inflation. Three pairs of in-sample regression and out of sample forecasting periods are chosen: 1951Q1 to 2008Q4 Vs 2009Q1 to 2010Q4; 1951Q1 to 2006Q4 Vs 2007Q1 to 2010Q4; 1951Q1 to 2004Q4 Vs 2005Q1 to 2010Q4, mean square forecasting error (MSFE) are calculated for each forecasting period. Similarly, we apply the same methodology to forecast Singapore GDP growth and inflation with the daily returns of Strait Times Index (STI) from January 1, 1986 to December 31,

2010^1, three forecasting periods are chosen: 2009Q1 to 2010Q4, 2007Q1 to 2010Q4 and 2005Q1 to 2010Q4.

4.1 Frequency Domain Filter

Ouliaris and Corbae (2005) proposes a new frequency domain filter for extracting the cyclical component of a time series from the level of a time series that easily handles stochastic and deterministic trends (and obviously, works for stationary series). Their approach yields a statistically consistent estimator of the ideal band pass filter. Using a series of monte carlo experiments, the authors show that their approach has much lower mean squared error than popular time domain filters (e.g. HP filter and BK filter) for a data generation process such as the growth rate of U.S. real output. Moreover, the proposed frequency domain filter has an important advantage over the Baxter and King (1999) and Hodrick-Prescott (1980) filters in that it does not require the investigator to set any parameters except the business cycle range. In our application, we extract components of GDP with range between 6 and 32 quarters (1.5 year to 8 years), or equivalently, 395 to 2088.5 days for the daily stock returns. The following figures show the results for filtered GDP, inflation and stock returns by applying the proposed frequency domain filter.

In the next section we shall apply the MIDAS methodology to aggregate both the filtered and unfiltered daily stock returns into the same frequency with GDP growth and inflation, implement forecasting and compare the predictive accuracy of different forecasting predictors.

---

[1] All the data in this thesis are downloaded from CEIC.

(a) US GDP

(b) Filtered US GDP

(c) Singapore GDP

(d) Filtered Singapore GDP
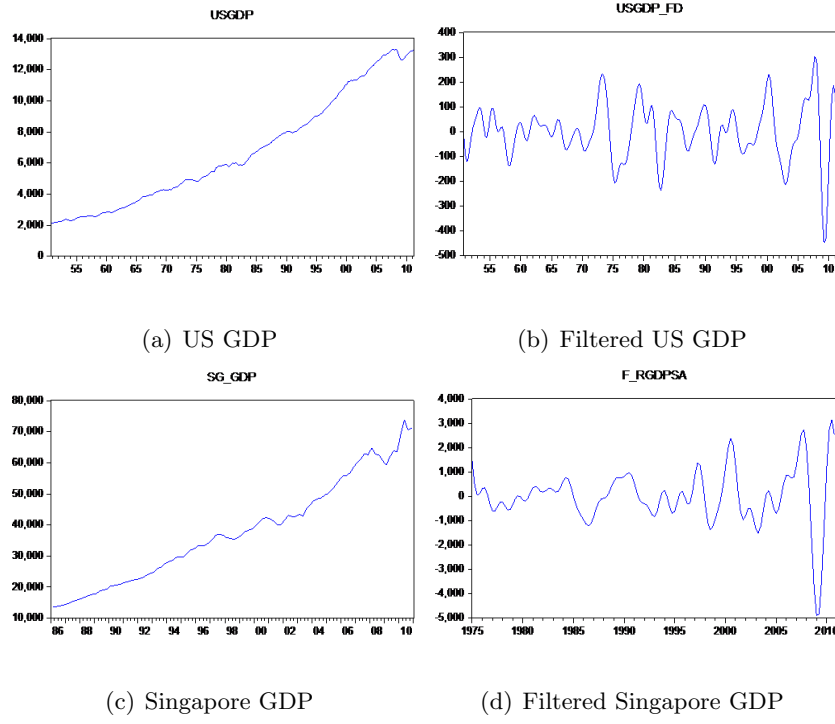
Figure 3: US(Singapore) GDP with filtered cycles from 1.5 years to 8 years

(a) US Inflation

(b) Filtered US Inflation

(c) Singapore Inflation

(d) Filtered Singapore Inflation
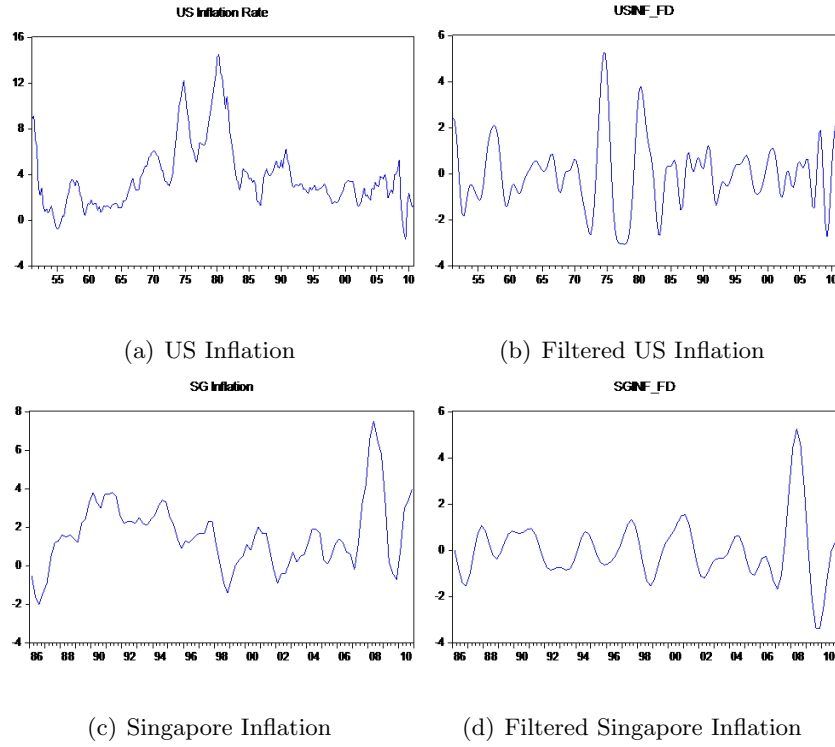
Figure 4: US(Singapore) Inflation with filtered cycles from 1.5 years to 8 years

12

(a) Dow Jones Index

(b) Filtered Dow Jones Index

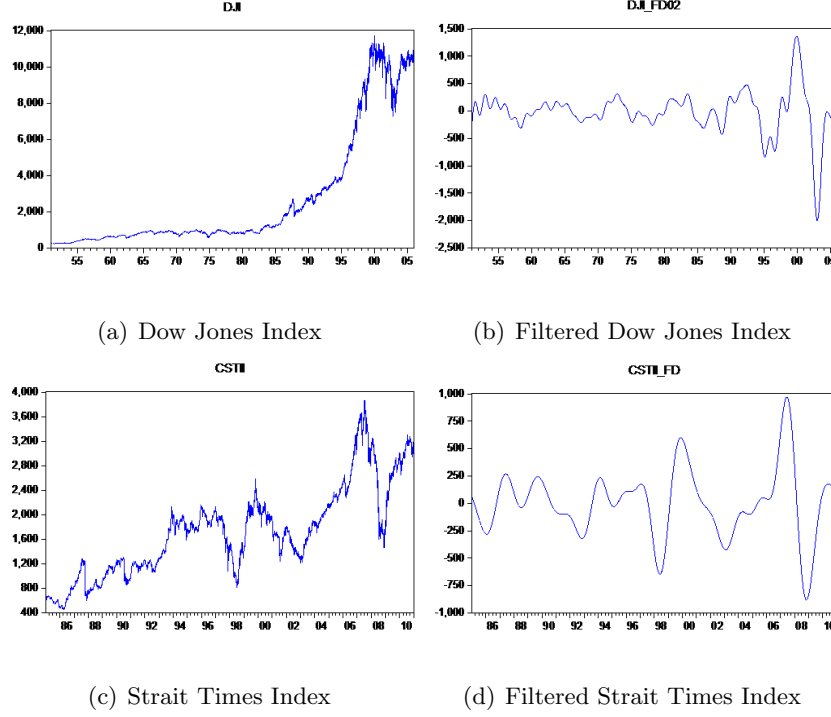(c) Strait Times Index

(d) Filtered Strait Times Index

Figure 5: DJI(STI) with filtered cycles from 395 days to 2088.5 days

4.2 Parameter Estimation

A crucial step in our MIDAS methodology is the specification of the parameters (θ_1,θ_2)

in the weighted aggregation functions (2) and (3). They not only determine the shapes of

these two aggregation functions, but also how many lags should be included in regression

(1). We intend to estimate (θ_1,θ_2) both in forecasting and nowcasting contexts. We propose

the following regression with an exponential Almon MIDAS weights (For the Beta MIDAS

weights, similar procedures apply):

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \beta_2 \sum_{k=0}^{K} B(k;\theta) L^{k/m} x_t^{(m)} + \varepsilon_t \tag{4}$$

Where,

$$B(k;\theta) = \frac{exp(\theta_1 k + \theta_2 k^2)}{\sum_{k=1}^{K} exp(\theta_1 k + \theta_2 k^2)}$$

13

Our objective here is to specify parameters β_0, β_1, β_2 and estimate θ_1, θ_2, by non-linear least square method. We first set reasonable initial values for θ_1 and θ_2, and aggregate the daily stock returns (unfiltered and filtered) into quarterly data by the Almon-weighted aggregation function. In the nowcasting context, regression (4) becomes like

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \beta_2 X_t + \varepsilon_t \tag{5}$$

where X_t represents the contemporary aggregated stock returns. We first estimate the values of $\beta_0, \beta_1, \beta_2$ in regression (5). Then by iterative calculations, we come out with the estimated values for all the five parameters (β_0, β_1, β_2) and (θ_1, θ_2). The results are presented in the following tables. [2]

Table 1. Parameter Estimation Using Filtered Daily

Stock Returns:Nowcasting

	US GDP	US Inflation	SG GDP	SG Inflation
$\hat{\beta}_0$	1.1456	-0.0272	251.3367	0.030
$\hat{\beta}_1$	0.8167	0.3000	1.0086	0.4519
$\hat{\beta}_2$	0.0409	0.0001	1.1851	0.0011
$\hat{\theta}_1$	4.1002	4.1035	0.1291	28.7821
$\hat{\theta}_2$	-33.5486	-30.1574	-27.5759	-0.2923

Initial values: $\theta_1 = 0.01$, $\theta_2 = -0.001$.

---

[2]Ghysels et al. (2005) reported the estimated parameters of the Almon MIDAS weight function in ICAPM (Intertemporal Capital Asset Pricing Model): $E_t[R_{t+1}] = \mu + \gamma Var_t[R_{t+1}]$. They showed that ($\theta_1 = -5.141$, $\theta_2 = -10.580$) for the full sample, and ($\theta_1 = -0.909$, $\theta_2 = -10.807$) and ($\theta_1 = -6.336$, $\theta_2 = -18.586$) for two subsamples respectively. They found that, in all three cases, the weights are a slowly declining function of the lag length. Please refer to the paper for details.

Table 2. Parameter Estimation Using Unfiltered Daily

Stock Returns:Nowcasting

	US GDP	US Inflation	SG GDP	SG Inflation
$\hat{\beta}_0$	1.5728	-0.00286	254.2632	0.0562
$\hat{\beta}_1$	0.9349	0.3211	1.0107	0.5108
$\hat{\beta}_2$	-1.2026	-0.0073	32.4763	-0.0206
$\hat{\theta}_1$	0.1304	3.1700	24.3923	14.7105
$\hat{\theta}_2$	-0.0017	-0.0939	-0.2100	-0.1508

Initial values: $\theta_1 = 0.01$, $\theta_2 = -0.001$.

Figure (6) and (7) show the relationship between the Almon weights and number of lagged days in the context of nowcasting, after inserting the estimated values($\hat{\theta}_1$ and $\hat{\theta}_2$) into the Almon weighted function.

As is shown in figure (6), with filtered stock returns, our results tell that only current day's stock returns information provides useful information for predicting US (Singapore) GDP growth and US inflation, with exception for Singapore inflation, which puts greatest weights around the 50th lagged day, and decline gradually both before and after that day, eventually die out before the 45th day or after 55th days. The former phenomenon,to some extent, coincides with the classical Efficient-Market Hypothesis, which states that current prices on traded assets (e.g., stocks, bonds, or property) already reflect all past publicly available information.

Figure (7) shows the Almon weights with estimated values of($\hat{\theta}_1$ and $\hat{\theta}_2$)from unfiltered stock returns. The Almon function shows good hump shape for US GDP growth, with highest portion of weight around the 50th lagged days, and decline gradually until the 100th days. For US inflation, the estimated Almon function reaches its highest point around the 20th lagged days, with the maximum value about 0.17, and becomes zero before the 10th
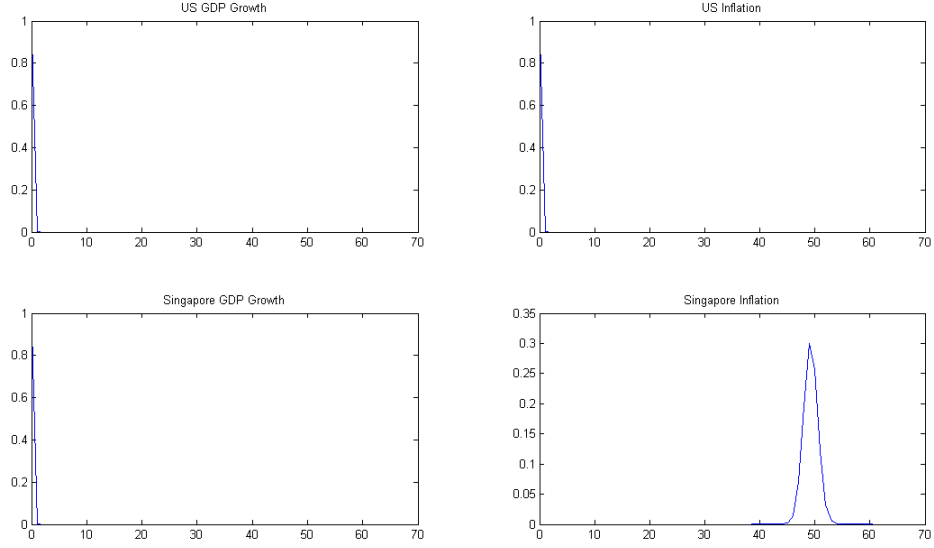
Figure 6: Almon Weights for Filtered Stock Returns-Nowcasting

lagged day or after 25 lagged days. Moreover, the estimated Almon function reaches its maximum value around the 60th lagged day for Singapore GDP growth (with a maximum value of 0.25), and the 50th lagged day for Singapore inflation (with a maximum value of 0.225).

In the context of forecasting, regression (4) becomes like

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \beta_2 X_{t-1} + \varepsilon_t \tag{6}$$

where X_{t-1} represents the lagged aggregated stock returns. We apply the same procedure as in the context of nowcasting, and present the estimated values both for the filtered and unfiltered stock returns in table 3 and table 4 respectively.

Figure (8) shows similar results in the context of forecasting for the US (Singapore) GDP growth and US inflation, but shows something odd for the Singapore inflation: Contrary to
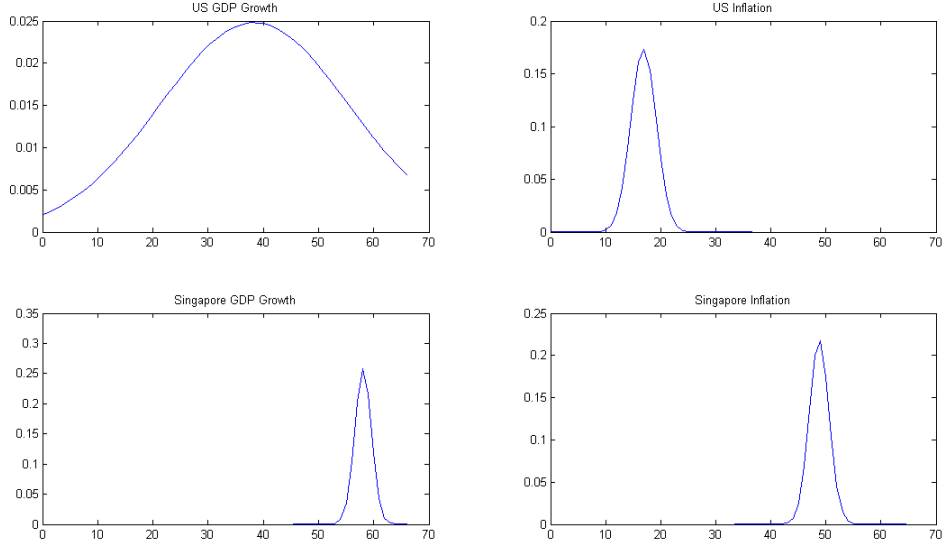
16

Figure 7: Almon Weights for Unfiltered Stock Returns-Nowcasting

a declining or hump-shaped Almon function, it is ever increasing with the number of lagged days, always putting the highest values on the farthest lagged days. However, we only show the situation with 30 lagged days here.

Table (4) shows the result with unfiltered stock returns, and figure (9) presents the corresponding Almon functions for each pair of estimated values. For US GDP growth, its Almon aggregated function reaches its maximum value (slightly below 0.025) between the lagged 50 and 100 days, and dies out gradually after that. For US inflation, the Almon function reaches its highest value between the lagged 50 and 60 days, and turns out to be zero either before the lagged 45th day or after the lagged 62th day. Similarly, the estimated Almon function for the Singapore GDP growth reaches its highest value 0.033 around the 40th lagged day, and becomes zero after 100 lagged days. For Singapore inflation, its Almon weights has the highest value of 0.1 around the 50th lagged day, and turns out to be zero

before the lagged 40th day or after the lagged 60th day.

Table 3. Parameter Estimation Using Filtered Daily

Stock Returns:Forecasting

	US GDP	US Inflation	SG GDP	SG Inflation
$\hat{\beta}_0$	0.0049	0.1233	0.0155	0.2387
$\hat{\beta}_1$	0.3608	0.9593	0.0715	0.8754
$\hat{\beta}_2$	-0.0000	-0.0000	0.0000	0.0015
$\hat{\theta}_1$	6.3313	35.3327	2.2267	0.4938
$\hat{\theta}_2$	-29.8455	-57.9604	-31.2831	-0.0000

Initial values: $\theta_1 = 0.01$, $\theta_2 = -0.001$.

Table 4. Parameter Estimation Using Unfiltered Daily

Stock Returns:Forecasting

	US GDP	US Inflation	SG GDP	SG Inflation
$\hat{\beta}_0$	0.0047	0.1006	0.0140	0.2756
$\hat{\beta}_1$	0.3680	0.9621	0.1343	0.8716
$\hat{\beta}_2$	0.2880	30.7183	2.7934	-57.2243
$\hat{\theta}_1$	9.7701	0.2450	0.2556	2.9783
$\hat{\theta}_2$	-0.0896	-0.0019	-0.0034	-0.0315

Initial values: $\theta_1 = 0.01$, $\theta_2 = -0.001$.
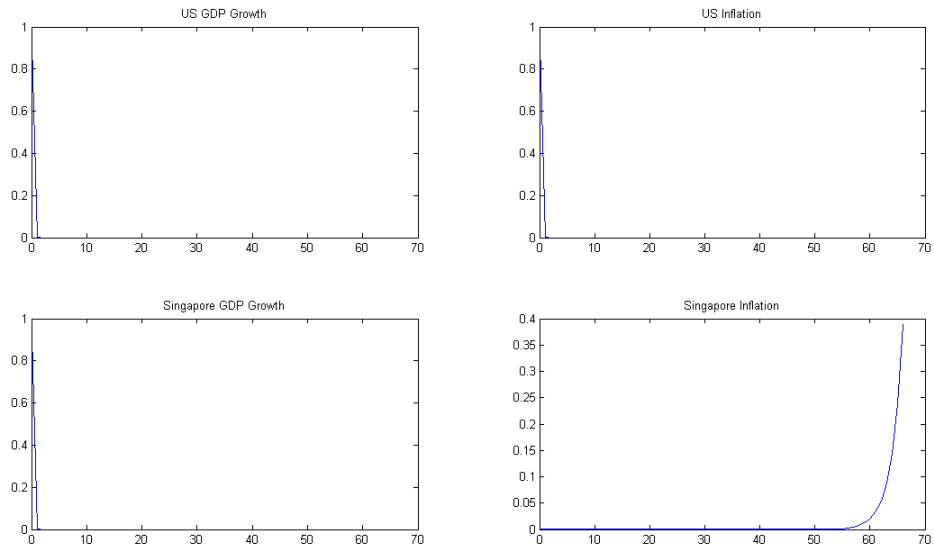
Figure 8: Almon Weights for Filtered Stock Returns-Forecasting
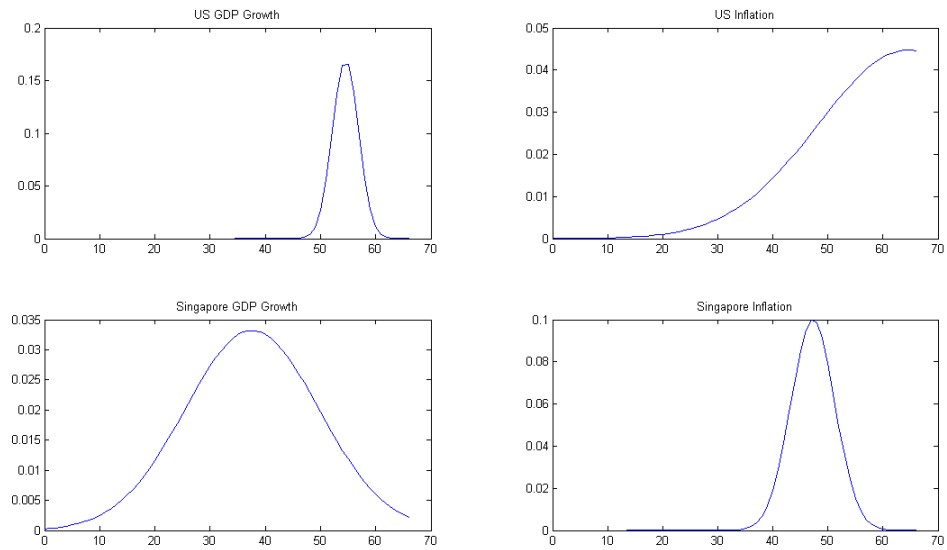

Figure 9: Almon Weights for Unfiltered Stock Returns-Forecasting

5 Forecasting Analysis

We consider the following forecasting models:

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \beta_2 S_t^q + \varepsilon_t \qquad (7)$$

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \beta_2 S_{t-1}^q + \varepsilon_t \qquad (8)$$

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \beta_2 s_t^{ud} + \varepsilon_t \qquad (9)$$

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \beta_2 s_{t-1}^{ud} + \varepsilon_t \qquad (10)$$

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \beta_2 s_t^{fd} + \varepsilon_t \qquad (11)$$

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \beta_2 s_{t-1}^{fd} + \varepsilon_t \qquad (12)$$

where Y_t represents the log-differenced unfiltered GDP (or inflation) for US and Singapore, $S_t^q (S_{t-1}^q)$ represents log-differenced unfiltered quarterly stock returns, $s_t^{ud}(s_{t-1}^{ud})$ represents the aggregated log-differenced unfiltered daily stock returns, and $s_t^{fd}(s_{t-1}^{fd})$ refers to the aggregated filtered daily stock returns. We do our forecasting both in the nowcasting and forecasting contexts, and in each context we choose the quarterly unfiltered stock returns as our benchmark models (e.g., (7) for nowcasting and (8) for forecasting, respectively). The purpose for doing this is because the quarterly stock returns data is at the same frequency with the GDP growth and inflation data, hence we can do forecasting straightforwardly, and by comparing the daily stock returns with the quarterly stock returns, we can figure out whether the high frequency stock returns data contain any useful information for forecasting GDP growth and inflation that the quarterly data don't grasp, it can also test the validity of our proposed MIDAS methodology in that whether it keeps as much useful information as possible while aggregating the data in a parsimonious way. Furthermore, we also want to figure out that, after filtering out the super-high frequency noise and possible trends from

the daily stock returns, whether our filtered daily stock returns would forecast better than the original data, or not.

We present our forecasting evaluation results for both the GDP growth and inflation in the following two tables. Mean Square Forecasting Errors (MSFE) for each predictor are calculated both in the nowcasting and forecasting contexts, and divided by the MSFEs of the benchmark models in each context respectively.

5.1 Forecasting Evaluation-GDP Growth

Table 5 US and Singapore GDP Growth Nowcasting & Forecasting

Forecasting Period	US GDP Growth Nowcasting		SG GDP Growth Nowcasting	
	s_t^{ud}	s_t^{fd}	s_t^{ud}	s_t^{fd}
2009Q1 to 2010Q4	2.787695731	1.113035133	0.569613448	0.934890251
2007Q1 to 2010Q4	1.270814403	1.336780972	0.857746753	0.917515765
2005Q1 to 2010Q4	1.237694582	1.177303433	0.866401599	0.917596546
	US GDP Growth Forecasting		SG GDP Growth Forecasting	
Forecasting Period	s_{t-1}^{ud}	s_{t-1}^{fd}	s_{t-1}^{ud}	s_{t-1}^{fd}
2009Q1 to 2010Q4	1.0678214	1.2072552	1.034018762	1.088790549
2007Q1 to 2010Q4	1.1483504	1.1146641	1.046224830	1.085246168
2005Q1 to 2010Q4	1.1002330	1.1049947	1.037264585	1.074520526

As is shown in table 5, the high frequency daily stock returns (both the filtered and unfiltered stock returns) help us little for forecasting US GDP growth, with calculated MSFEs (Mean Squared Forecasting Error) strictly higher than 1 compared to the benchmark models, and we can not easily distinguish the superiority of these two predictors both in the nowcasting and forecasting contexts. For instance, for the forecasting period: 2009Q1 to 2010Q4, the relative MSFE of the unfiltered stock returns is much larger than the filtered

stock returns, however, during the forecasting period from 2007Q1 to 2010Q4, the relative MSFE of the filtered stock returns is greater than the unfiltered stock returns. Finally, for the forecasting period from 2005Q1 to 2010Q4, the relative MSFE of the unfiltered stock returns is, again, greater than the filtered stock returns. In the forecasting context, on the other hand, the situation is reversed. For example, during the forecasting period from 2009Q1 to 2010Q4, the relative MSFE of the filtered stock returns is greater than the unfiltered stock returns, while during the period for 2007Q1 to 2010Q4, the unfiltered stock returns forecast than the filtered stock returns. Finally, during the forecasting period from 2005Q1 to 2010Q4, the relative MSFE of the unfiltered stock returns is sightly greater than the unfiltered stock returns.[3]

For the Singapore GDP growth, we see some positive signs with our MIDAS methodology in the nowcasting context, in that both the filtered and unfiltered stock returns improve our predictions. Particularly for the unfiltered stock returns, which improves the relative MSFE by 43% for the forecasting periods of 2009Q1 to 2010Q4 and 15% for both the forecasting periods of 2007Q1 to 2010Q4 and 2005Q1 to 2010Q4. As for the filtered stock returns, moderate improvements also have been made, from roughly 6.5% for the forecasting period 2009Q1 to 2010Q4 to 8% for the both the forecasting periods of 2007Q1 to 2010Q4 and 2005Q1 to 2010Q4. In the forecasting context, on the other hand, we see little improvement with our filtered and unfiltered high frequency stock returns data. One possible explanation is that our filtering process may discard much useful information contained in the noise or

---

[3]Clements and Galvao (2008) also applied the MIDAS approach to forecast the US output growth, but their method was different from ours in that they included an autoregressive term in the regression. They showed that the MIDAS approach with an AR term reduces RMSE sizeably compared to AR model when monthly data on industrial production and capacity utilization are available on the current quarter.

trend that are important for predicting Singapore GDP growth. Moreover, as is easily seen from the table, the unfiltered stock returns forecast better than the filtered stock returns during all the three forecasting periods. For this reason, we can safely conclude that the unfiltered stock returns forecast consistently better than the filtered stock returns with respect to the Singapore GDP growth forecasting.

5.2   Forecasting Evaluation-Inflation

Similarly, Y_t represents the quarterly inflation level (US and Singapore), $s_t^{ud}(s_{t-1}^{ud})$ represents the log-differenced unfiltered daily stock returns, and $s_t^{fd}(s_{t-1}^{fd})$ refers to the filtered daily stock returns. We present our forecasting evaluation results in the following tables.

Table 6 US and Singapore Inflation Nowcasting & Forecasting

| Forecasting Period | US Inflation Nowcasting | | SG Inflation Nowcasting | |
	s_t^{ud}	s_t^{fd}	s_t^{ud}	s_t^{fd}
2009Q1 to 2010Q4	0.777846704	0.79398270	0.82592301	0.701006919
2007Q1 to 2010Q4	0.863307462	1.044959807	0.92565959	0.762610611
2005Q1 to 2010Q4	0.883402504	0.977297263	0.92127627	0.772437858
Forecasting Period	US Inflation Forecasting		SG Inflation Forecasting	
	s_{t-1}^{ud}	s_{t-1}^{fd}	s_{t-1}^{ud}	s_{t-1}^{fd}
2009Q1 to 2010Q4	0.986398286	1.012689699	0.832908211	0.520009528
2007Q1 to 2010Q4	0.953388667	0.946305596	0.913009788	0.593589869
2005Q1 to 2010Q4	0.951391058	0.953525680	0.908407553	0.619014245

Table 6 contains the main findings of this thesis, therefore we will discuss them in more detail here. Our results show that the high frequency stock returns data (both the filtered and unfiltered stock returns) do contain more useful information than the quarterly data

in forecasting inflation. In the context of nowcasting, the unfiltered stock returns forecast better than the filtered stock returns for US inflation. For instance, for the forecasting period from 2009Q1 to 2010Q4, the relative MSFE of the unfiltered stock returns is slightly lower than the filtered stock returns; for the period of 2007Q1 to 2010Q4, the filtered stock returns forecast worse than the benchmark, while the unfiltered stock returns improve the MSFE by roughly 15%; for the forecasting period of 2005Q1 to 2010Q4, the filtered stock returns improve the MSFE by mere 2%, while the the unfiltered stock returns improve the MSFE by roughly 12%.

However, for Singapore inflation, the filtered daily stock returns forecast better than unfiltered stock returns for the all the three forecasting periods. The unfiltered stock returns improves the MSFE by 17%, while the filtered stock returns improves the MSFE by 30% during the forecasting period from 2009Q1 to 2010Q4; for the forecasting period of 2007Q1 to 2010Q4, the unfiltered stock returns reduce the MSFE (compared to the benchmark model) by 7%, while the filtered stock returns reduce the MSFE by 24%. Finally, during the forecasting period from 2005Q1 to 2010Q4, the unfiltered stock returns reduce the relative MSFE by 8%, while the filtered stock returns reduce the relative MSFE by 23%.

Furthermore, for US inflation, predictors in the nowcasting context perform better than in the forecasting context, this is easily to understand as in the nowcasting context, predictors contain more information than in the forecasting context. However, it is hard to tell which forecaster perform better in the forecasting context, as the relative MSFEs vary. During the period of 2009Q4 to 2010Q4, the relative MSFE of the filtered stock returns is greater than the unfiltered stock returns. However, for the period of 2007Q1 to 2010Q4, the relative MSFE of the unfiltered stock returns is greater than the filtered stock returns.

24

Finally, for the period of 2005Q1 to 2010Q4, the relative MSFE of the filtered stock returns is greater than the unfiltered stock returns.

Lastly, our forecasting result for the Singapore inflation seems some kind of abnormal, as we observe from the table 6 that both the unfiltered and filtered daily stock returns perform better in the forecasting context than in the nowcasting context, specially for the filtered daily stock returns data, with average 40% improvements compared to the benchmark model for all the three chosen forecasting periods in the forecasting context from rough 25% improvements in the nowcasting context. And it is easily observed that the filtered stock returns forecast straightly better than the unfiltered stock returns. Which affirms that, after filtering out the high frequency noise and low frequency trend, the remaining cycles do improve the predictive accuracy of our prediction for inflation.

5.3 Diebold-Mariano Test

Diebold and Mariano(1995) considers model-free tests of forecast accuracy that are directly applicable to quadratic loss functions, multi-period forecasts, and forecast error. The basic ideas are presented as follows.

Let y_t be a covariance stationary are ergodic process, e.g. an ARMA(p,q) process with Wold representation

$$
\begin{aligned}
y_t &= \mu + \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}, \varepsilon_t \sim WN(0, \sigma^2) \\
&= \mu + \varepsilon_t + \psi_1 \varepsilon_{t-1} + \psi_2 \varepsilon_{t-2} + \cdots
\end{aligned}
$$

and let $I_t = \{y_t, y_{t-1}, \cdots\}$ denote the information set available at time t. $y_{t+h|t}^1$ and $y_{t+h|t}^2$ denote two competing forecasting of y_{t+h} based on I_t. For example, $y_{t+h|t}^1$ could be computed from an AR(p) model and $y_{t+h|t}^2$ could be computed from an ARMA(p,q) model. The

forecast errors from the two models are

$$\varepsilon^1_{t+h|t} = y_{t+h} - y^1_{t+h|t}, \varepsilon^2_{t+h|t} = y_{t+h} - y^2_{t+h|t}$$

The h-step forecasts are assumed to be computed for $t = t_0, \ldots, T$ for the total of T_0 forecasts giving

$$\{\varepsilon^1_{t+h|t}\}^T_{t_0}, \{\varepsilon^2_{t+h|t}\}^T_{t_0}$$

Because the h-step forecasts use overlapping data the forecast errors in $\varepsilon^1_{t+h|t}{}^T_{t_0}$ and $\varepsilon^2_{t+h|t}{}^T_{t_0}$ will be serially correlated. The accuracy of each forecast is measured by a particular loss function

$$L(y_{t+h}, y^i_{t+h|t}) = L(\varepsilon^i_{t+h|t}), i = 1, 2$$

Some popular loss functions are

- Squared error loss: $L(\varepsilon^i_{t+h|t}) = (\varepsilon^i_{t+h|t})^2$

- Absolute error loss: $L(\varepsilon^i_{t+h|t}) = |\varepsilon^i_{t+h|t}|$

To determine if one model predicts better than another we may test null hypotheses

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

against the alternative

$$H_1 : E[L(\varepsilon^1_{t+h|t})] \neq E[L(\varepsilon^2_{t+h|t})]$$

The Diebold-Mariano test is based on the loss differential

$$d_t = L(\varepsilon^1_{t+h|t}) - L(\varepsilon^2_{t+h|t})$$

The null of equal predictive accuracy is then

$$H_0 : E[d_t] = 0$$

26

this can be done by regressing d_t on a constant and any $ARMA$ terms that may be suitable, and testing whether the constant term is significantly different from zero.

Herein we apply the Diebold-Mariano test to compare the predictive accuracy of various forecasters both in the nowcasting and forecasting contexts. We select the squared error loss as our loss function in our application. Moreover, we also do a cross comparison between predictors in the nowcasting and forecasting context. The results are shown in the following table.

Table 7 Diebold-Mariano Test

	US GDP Growth	US Inflation	SG GDP Growth	SG Inflation
Nowcasting	$d_t^{(gdp)}$	$d_t^{(inf)}$	$d_t^{(gdp)}$	$d_t^{(inf)}$
s_t^{ud}/S_t^q	0.0000116**	-0.188370^*	-0.000113	-0.135341^*
s_t^{fd}/S_t^q	0.00000866	-0.036678	-0.0000698	-0.391223^*
s_t^{fd}/s_t^{ud}	-0.00000295	0.151693^*	0.0000435	-0.255882^*
Forecasting	$d_{t-1}^{(gdp)}$	$d_{t-1}^{(inf)}$	$d_{t-1}^{(gdp)}$	$d_{t-1}^{(inf)}$
s_{t-1}^{ud}/S_{t-1}^q	0.00000680	-0.049238	0.000147^*	-0.014677
s_{t-1}^{fd}/S_{t-1}^q	0.00000676	-0.022845	0.0000583	-0.473673^{**}
$s_{t-1}^{fd}/s_{t-1}^{ud}$	-0.000000035	0.026393	-0.0000891	-0.458996^{**}
Cross Comparison	$d_{t/t-1}^{(gdp)}$	$d_{t/t-1}^{(inf)}$	$d_{t/t-1}^{(gdp)}$	$d_{t/t-1}^{(inf)}$
s_t^{ud}/s_{t-1}^{ud}	0.00000280	-0.063619	-0.000194^{***}	-0.072758
s_t^{fd}/s_{t-1}^{fd}	0.00000572	0.061681	-0.0000618	0.130356^{**}
S_t^q/S_{t-1}^q	-0.00000761^*	0.075513	0.0000663	0.047906

$*,**$ and $***$: level of significance at 10%, 5% and 1% respectively.

As is shown in the upper panel of table 7. In the nowcasting context, the forecasting power of the aggregated unfiltered daily stock returns is weaker than the benchmark model for the US GDP growth(with the quarterly stock returns as regressor) at a significant

level of 5%, while there are no significant differences between aggregated filtered daily stock returns and aggregated unfiltered daily stock returns. For US inflation, however, the aggregated unfiltered daily stock returns works better than both the aggregated filtered daily stock returns and our benchmark predictor at a significant level of 10%, while there is no significant difference between the aggregated the filtered daily stock returns and quarterly stock returns. Moreover, there are no significant predictive power differences among all our three forecasters for the Singapore GDP growth. Finally, our results for the Singapore inflation show something new: The application of the MIDAS methodology greatly improves forecasting accuracy compared to the benchmark predictor. Particularly, the aggregated filtered daily stock returns forecast better than aggregated unfiltered daily stock returns, which illustrates that, after filtering out the trends and high frequency noise, the remaining stock returns data still contain a lot of useful information.

In the forecasting context, the result shows there are no significant predictive differences among all the three forecasters both for the US GDP growth and US inflation, e.g.,with d_t very close to zero and p-value greater than 10%. For Singapore GDP growth, the test shows that the unfiltered daily stock returns forecast slightly worse than the benchmark predictor, while there are no significant predictive difference between filtered stock returns and the benchmark predictor, or between the unfiltered and filtered stock returns. For Singapore inflation, the filtered stock returns forecast better than both the benchmark predictor and the unfiltered stock returns at a significance level of 5%. Meanwhile, there are no significant difference between the unfiltered stock returns and the benchmark predictor.

The cross comparison between predictors in the nowcasting context and forecasting context is shown in the bottom panel of table 7, our test shows that, for US GDP growth, there

28

are no significant differences between the contemporary unfiltered daily stock returns and its lagged counterpart, or between the contemporary filtered daily stock returns and its lagged counterpart, which tells new information gained in current quarter does not improve the predictive accuracy of these two forecasters. However, for the benchmark predictors, the inclusion of new information in current quarter indeed helps improve our prediction, though at a very weak level. Secondly, there are no significant predictive differences among all the three contemporary predictors and their lagged counterparts for the US inflation forecasting. As for Singapore GDP growth, the contemporary unfiltered daily stock returns works better than its lagged counterpart at a significance level of 1%, but there are no significant differences for the other two predictors. Finally, for Singapore inflation, the contemporary filtered daily stock returns forecast even worse than its lagged counterpart, which means that, the inclusion of the new information makes the prediction for inflation even worse. It is contrary to reality and probably due to some errors made in the process of data interpretation and our forecasting methodology. As for the two other predictors, we again see little differences between the contemporary predictors and their lagged counterparts.

6   Conclusion

In this thesis we study the predictive power of daily stock returns on GDP growth and inflation. We filter the daily stock returns with a newly proposed frequency domain filter, and aggregate the daily data with MIDAS functions using estimated parameter values. We find that predictors with MIDAS regressions generally perform better in inflation forecasting than in GDP growth forecasting. Specifically, our forecasting with MIDAS regressions improve greatly for Singapore inflation compared to the the benchmark model, both with

unfiltered stock returns and filtered stock returns. Moreover, the predictors with filtered stock returns performs better than the predictors with unfiltered stock returns, and predictor with lagged filtered stock returns perform better than contemporary filtered stock returns(with an average improvement of 45% versus 25% in relative MSFE with respect to the benchmark model). For US inflation, however, predictors with unfiltered stock returns perform better with filtered stock returns, and predictor with contemporary stock returns works better than with lagged stock returns. For Singapore GDP growth, we find that only contemporary stock returns have more forecasting power than the benchmark model, and like US inflation, unfiltered daily stock returns forecast better than filtered daily stock returns. For US GDP growth, on the other hand, we see little improvement with our MIDAS regressions.

References

[1] Alper, C.E., Fendoglu, S. and Saltoglu, B. (2008). Forecasting Stock Market Volatilities Using MIDAS Regressions: An Application to the Emerging Markets. *Unpublished Paper.*

[2] Alper, C.E., Fendoglu, S. and Saltoglu, B. (2011). MIDAS Volatility Forecast Performance Under Market Stress: Evidence from Emerging and Developed Stock Markets. *Working Paper.*

[3] Andreou, E., Ghysels, E. and Kourtellos, A. (2009). Should Macroeconomic Forecasters Use Daily Financial Data?. *UCY Working Paper.*

[4] Andreou, E., Ghysels, E. and Kourtellos, A. (2010). Regression Models with Mixed Sampling Frequencies. *Journal of Econometrics* **158(2)**, pp.246-261.

[5] Clements, M.P. and Galvo, A.B. (2008). Macroeconomic Forecasting with Mixed-Frequency Data:Forecasting Output Growth in the United States. *Journal of Business and Economic Statistics* **26(4)**, pp.546-54.

[6] Clements, M.P. and Galvo, A.B. (2009). Forecasting US Output Growth Using Leading Indicators: An Appraisal Using MIDAS Models. *Journal of Applied Econometrics* **24(7)**,pp. 1187-1206.

[7] Galvo, A.B. (2011). Changes in Predictive Ability with Mixed Frequency Data. *Working Paper, Queen Mary University of London.*

[8] Ghysels, E., Santa-Clara, P. and Valkanov, R. (2004). The MIDAS Touch: Mixed Data Sampling Regression Models. *Working Paper.*

[9] Ghysels, E., Santa-Clara, P. and Valkanov, R. (2005). There Is a Risk-Return Trade-off After All. *Journal of Financial Economics* **76(3)**, pp.509-48.

[10] Ghysels, E., Santa-Clara, P. and Valkanov, R. (2006). Predicting Volatility: Getting the Most Out of Return Data Sampled at Different Frequencies. *Journal of Econometrics* **131(1/2)**, pp.59-95.

[11] Ghysels, E., Santa-Clara, P. and Valkanov, R. (2007). MIDAS Regressions: Further Results and New Directions. *Econometric Reviews* **26(1)**, pp.53-90.

[12] Kuzin, V., Marcellino, M. and Schumacher, C. (2011). MIDAS versus Mixed-Frequency VAR: Nowcasting GDP in the Euro Area. *International Journal of Forecasting* **27(2)**, pp.529-542.

[13] Ouliaris, S. and Corbae, P.D. (2005). Extracting Cycles from Non-stationary Data. *Frontiers of Analysis and Applied Research: Essays in honor of Peter C. B. Phillips.* Cambridge University Press: London.

[14] Tay, A.S. (2006). Mixing Frequencies: Stock Returns as a Predictor of Real Output Growth. *Working Paper, Singapore Management University.*

[15] West, K. (2006). Forecast Evaluation. *Handbook of Economic Forecasting* **1**, pp.99-134.