Forecasting the Price of Gold Using Dynamic Model Averaging

Goodness C. Aye
University of Pretoria
Rangan Gupta
University of Pretoria
Shawkat Hammoudeh
Drexel University
Won Joong Kim
Konkuk University

_____

Department of Economics
University of Pretoria
0002, Pretoria
South Africa
Tel: +27 12 420 2413

Forecasting the Price of Gold Using Dynamic Model Averaging

Goodness Aye[a], Rangan Gupta[a], Shawkat Hammoudeh[b], Won Joong Kim[c]

Abstract

We develop models for examining possible predictors of the return on gold that embrace six global factors (business cycle, nominal, interest rate, commodity, exchange rate and stock price factors) and two uncertainty indices (the Kansas City Fed's financial stress index and the U.S. Economic uncertainty index). Specifically, by comparing with other alternative models, we show that the dynamic model averaging (DMA) and dynamic model selection (DMS) models outperform not only a linear model (such as random walk) but also the Bayesian model averaging (BMA) model for examining possible predictors of the return of gold. The DMS is the best overall across all forecast horizons. Generally, all the predictors show strong predictive power at one time or another though at varying magnitudes, while the exchange rate factor and the Kansas City Fed's financial stress index appear to be strong at almost all horizons and sub-periods. However, the forecasting prowess of the exchange rate is supreme.

[a] Department of Economics, University of Pretoria, Pretoria, South Africa.
Email: rangan.gupta@up.ac.za.
[b] Lebow College of Business, Drexel University, Philadelphia, United States and IPAG Business School, Paris, France.
Email: hammousm@drexel.edu.
[c] Corresponding author. Department of Economics, Konkuk University, Seoul, Korea.
Email: wjkim72@konkuk.ac.kr.

1. Introduction

Gold serves several functions in the world economy, and its links with financial and macroeconomic variables are quite well-established (Pierdzioch et al., 2014a, b). It has a monetary value and is sought after by central banks to be part of their international reserves which fulfil many purposes (Gupta et al., 2014).[1] It has industrial uses and can be transformed into jewelry. In modern finance, it is used as a hedge against inflation and a safe haven during crises. Gold has also other distinguished characteristics. Its supply is accumulated over the years and its global annual physical production can be as small as 2% of total supply, thereby in contrast to other commodities its annual production may not sway its price as other factors do. Moreover, unlike prices of bonds, the gold price depends on future supply and demand, and thus it is forward-looking.

A myriad of other global factors affect the price of gold (Pierdzioch et al., 2014a), which in turn, can be primarily classified into six categories: The business cycle factor; the nominal factor; the interest rate factor; the commodity factor; the exchange rate market factor; and the stock price factor. As is standard in the literature, one factor is created based on the first principal component of each set or bloc of variables. Thus, for example, for the business cycle factor, we use the first principle component of the four stationary and standardized variables. All the factors are latent, but they still capture the common component of the variables in the group. We add to these six factors, the Kansas City Fed's financial stress index and the U.S. economic uncertainty which can impact the gold price returns because gold is used as a safe haven and those variables deal with fear and anxiety (Ciner et al., 2013). The business cycle factor (Pierdzioch et al., 2014b) is represented by an economic activity variable, which is the industrial production of the first and second largest economies, namely the United States and China, as

---

[1] On an average year, more than 62% of gold supply is used in jewelry, more than 18% is used in industrial applications and about 18% is held in central banks' foreign reserves.

well as for Japan and Europe. An improvement in this economic activity will increase the industrial and jewelry demands for gold as a result of increased world production and income.

The nominal factor refers to inflation as represented by the CPI for the United States, EU, China and Japan and the money supply growth for the United States and EU. As indicated, gold is a hedge against inflation. Moreover, gold prices are sensitive to changes in interest rates which represent the opportunity cost of hoarding gold. This factor includes the interest rates for the U.S. and EU, and LIBOR. Gold is also a commodity that is affected by movement in commodity prices. When there is a commodity boom, gold co-moves in sympathy because there are common macroeconomic variables that comove related and unelated commodities (Pindyck and Rotemberg, 1990). The gold price return is also influenced by changes in the foreign exchange markets since gold is considered a currency (Sjaastad, 2008; Sjaastad and Scacciavillani, 1997). This relationship is theoretically explained by the law of one price. A weakness in the dollar exchange rate will motivate investors and traders to seek the safety of gold (Ismail et al., 2009). Finally, gold is an asset class that is related to financial asset classes like the stocks. Gold and stocks may move together but they may depart depending on the business cycle.

Methodologically against this backdrop, this paper applies the dynamic model averaging (DMA) strategy developed by Raftery *et al.* (2010), which can also be used for the dynamic model selection (DMS) methodology. In this mythology, a single but potentially different model can be employed as a forecasting model at each point of time to forecast the returns on gold based on the wide array of predictors discussed above, which may move in sync or in different directions. The DMA or DMS seems ideally suited for forecasting a variable as volatile as the returns on gold (see Figure 1), which have gone through periods of major events and crises. These methodologies not only allow the forecasting model to change over time, but also the coefficients in each model to evolve over time simultaneously. These two methods involve standard econometric methods used for state space models, such as the Kalman filter, but through some sensible empirical approximations, they achieve vast gains in computational

efficiency so as to allow one to apply DMA and DMS despite the computational difficulties. The advantages of using the DMA approach over standard predictive regression methods are three folds. (1) The coefficients on the predictors can change over time, while the conventional use of dummy variables to capture possible structural breaks or recursive methods to take care of time variations is not suited to solve this issue. Hence ideally, one needs to build models designed to capture the changing of slope parameters of the models. (2) Given that we use a wide variety of predictors, the number of potential predictors is in general quite large, thereby in the process making the number of models quite large as well. For instance, if the set of models is defined by whether each of the m potential predictors are included or excluded, then we have 2^m possible models to work with. This would raise substantial statistical problems involved in model selection, which in turn, has resulted in researchers resorting to the Bayesian model averaging (BMA) as in Wright (2008) and Tortora (2009). (3) Finally, the model that is relevant for forecasting can itself potentially change over time due to a change in the set of predictors or with some variables better capable of forecasting during, for example, recessions than expansions. This makes the situation even more complicated, since with 2^m models the number of combinations of models required to be estimated in order to forecast at time τ is $2^{m\tau}$ to allow for the possibility of a different model being applicable at each point in time. Even in relatively simple forecasting exercises, it can be computationally impossible to forecast by simply going through all of these $2^{m\tau}$ combinations.

To the best of our knowledge, this is the first study to forecast returns on gold using time-variations in parameters as well as models. Based on monthly data on the predictors and using an out-of-sample period of 1999:10-2013:11 (with an in-sample of 1992:02-1999:09), we compute multi-steps-ahead forecasts for gold returns based on recursive estimation of the models over the out-of-sample period. Note that the start and end-point of our entire sample is purely driven by

data availability, while the starting point of out-of-sample corresponds to the period when the returns on gold reached its peak (see Figure 1).

The results show that is clear among all the models considered, namely the time-varying parameter (TVP), DMA, DMS, Bayesian model averaging (BMA) and the benchmark random-walk (RW) models, the DMA and DMS models forecast generally well, with DMS being the best overall. These results are supported by both diagnostics, namely the sum of the log predictive likelihoods and the mean squared forecast error (MSFE). It can therefore be concluded that allowing for model uncertainty and not only parameter uncertainty would improve the forecasting performance of these models. One major advantage of the DMA is that the framework is capable of identifying which factors play important roles at each point in time in forecasting gold returns over different horizons of the out-of-sample period, based on time-varying posterior inclusion probabilities. The results show that the good predictors of the gold prices are the exchange rate, the financial stress index, the interest rate and the business cycle factors which can be considered as the leaders of the forecasting procession over the forecast periods. Nevertheless, the forecasting prowess of the exchange rate is supreme.

There are benefits from finding the right model that forecasts the gold price returns more accurately than others. The out-of-sample forecasting offers informational advantage for monetary policymakers, hedge fund managers and international portfolio managers which can be used in gauging future inflation, estimating demand for jewellery, discerning investment in precious metals and other commodities and assessing the future movement of the dollar exchange rate. It also provides additional support to the hypothesis that the economic activity, foreign exchange, and stocks, interest rates and gold prices are driven by the same information sets of predictors. This lends support to the belief that the world gold market is currently dominated by the U.S. dollar bloc.

The reminder of the paper is organized as follows. Section 2 presents a brief of review of the literature. Section 3 describes the methodology and Section 4 discusses the results. Section 5 concludes the paper.

2. Review of related literature

The literature on forecasting changes in the gold returns is not as generous as it is on gold return movement and volatility, and the methods employed in forecasting are not as sophisticated either. Ismail et al. (2009) use the multiple linear regression to predict the future gold prices using a set of economic factors selected based on "a hunch of experts" including Commodity Research Bureau future index, U.S. dollar to euro foreign exchange rate, inflation rate, money supply, S&P 500 index, Treasury Bill and U.S. nominal effective dollar index. The predictions are evaluated relative to the performance of a naïve (random-walk) model. The results show that the considered models achieve higher level of predictive accuracy compared to the naïve model.

Sjaasdat (2008) investigates the theoretical and empirical relationships between the price of gold and the major exchange rates, using forecast error data. The author finds the floating exchange rates among the major currencies to be a major source of price instability in the global gold market since the termination of the Bretton Woods international monetary system. Moreover, since this market is currently dominated by the U.S. dollar bloc, appreciations or depreciations of the dollar would have strong effects on the gold price expressed in other currencies. The author finds support for the hypothesis of market efficiency for the world gold market during the 1991–2004 period. However, he has not found that the major gold producers of the world (Australia, South Africa, and Russia) appear to have no significant influence over the world gold price.

Shafiee and Topal (2010) investigate the relationship between the gold price and key influencing factors such as the oil price and global inflation over the last 40 years. The authors employ the trend reverting jump and dip diffusion model to forecasting natural-resource commodity prices including the price of gold. The model examines the fluctuations of prices in terms of a long-term trend reversion component, a diffusion component and a jump or dip component. The model is then applied to forecast the gold price for the next 10 years. The forecast results show that the gold price is predicted to stay abnormally high up to the end of 2014, after that it would revert to the long-term trend until 2018. Lili and Chengmei (2013) use the Factor-Augmented Vector Auto-regression (FAVAR) model to extract three common factors from the panel data series, representing financial market indices, gold reserves and energy prices, and macroeconomic indicators that affect gold prices.

Pierdzioch et al. (2014a) employ a real-time (recursive, as also in our case) forecasting approach to examine whether publicly available information on a large set of financial and macroeconomic variables would help to forecast out-of-sample monthly excess returns of gold under uncertainty. This approach also accounts for the possibility that the optimal forecasting model may vary over time. The results show that using forecasts implied by this forecasting approach to set up a simple trading rule does not necessarily lead to a superior performance relative to a buy-and-hold strategy. Pierdzioch et al. (2014b) also find some evidence that the international business cycle has predictive power for gold-price fluctuations. They suggest a behavioral-finance approach to study the quality of out-of-sample forecasts from the perspective of forecasters with potentially asymmetric loss functions.

Baur and Glover (2014) develop and test empirically several models incorporating heterogeneous expectations of agents, specifically fundamentalists and chartists, for the gold market. The empirical evidence shows that both agent types are important in explaining

historical gold prices. However, the 10-year bull run of gold in the early 2000s is consistent with the presence of agents extrapolating long-term trends.

On the benefits of forecasting the gold price, Apergis (2014) investigates whether the gold price has a reliable out-of-sample relationship with the Australian dollar/U.S. dollar nominal and real exchange rates using daily and quarterly data, respectively. Using an Error Correction Model (ECM), the empirical results suggest that the out-of-sample predictive ability is strong and robust across short- and long-run horizons, thus offering informational availability for monetary policymakers, hedge fund managers and international portfolio managers. They also provide additional support to the hypothesis that both markets are driven by the same information sets.

In summary, the papers that are most closely related to our paper is that of Pierdzioch et al. (2014a, b), in the sense that those authors look at similar set of predictors, though fewer in numbers and primarily concentrated for the U.S., and also allow for model uncertainty. However, unlike the papers by Pierdzioch et al. (2014a, b), we also model for nonlinearity both in- and out-of-sample, in the relationship between gold returns and its predictors by allowing for time-varying coefficients. The recursive (real-time) scheme adopted by Pierdzioch et al. (2014a, b) only allows for nonlinearity over the out-of-sample. And as we show, the constant parameter models, namely the BMA and a special case of DMA where the inclusion of the predictors varies over time only (but not the coefficients), based on recursive estimations over the out-of-sample is easily outperformed by the DMA and DMS. Further, we are able to provide information on the importance of specific predictors, for each forecasting horizon at each point of the out-of-sample based on posterior inclusion probabilities. This forecasting information matters and is helpful because it relates relevant predictors to specific time periods and saves investors from incurring losses and policy-makers from making mistakes at the wrong time by helping them obtain better signals.

3. Methodology

The models considered for forecasting the gold price return in this study include the time-varying parameter (TVP) model, the dynamic model averaging (DMA) model, the dynamic model selection (DMS) model, the Bayesian model averaging (BMA) model and a random walk (RW) model.

The time-varying parameter (TVP) models employ state space methods such as the Kalman filter, which is commonly used in empirical macroeconomic research on structural analysis and forecasting. These types of models, however, do not allow predictors to vary over time (Koop and Korobilis, 2012).[2] If large sets of predictors are used, then the TVP models tend to over-fit in-sample, and therefore have a poor out-of-sample forecasting performance. Even extensions of these models such as the TVP-VAR models suffer from this same limitation (Koop and Korobilis, 2012). To address these shortcomings in the TVP models, the DMA models present a possible and better alternative.

The dynamic model averaging (DMA) model, as the name implies, averages across various models. Moreover, BMA which is a specific form of the DMA model is only used for static linear models with parameter uncertainty. The uncertainty is accounted for in the BMA by averaging over all the sets of possible explanatory variables that may be included in the models (Raftery *et al.*, 2010). A shortcoming of the BMA approach is that it is limited to static models only. It was observed that the dynamics of the various models tend to follow a hidden Markov chain which can be incorporated in DMA using a recursive updating method such as the Kalman filter (Raftery *et al.*, 2010). Using the DMA framework, the BMA can be easily derived by simply excluding any dynamics from the DMA estimates.

To understand the econometric methodology regarding DMA, suppose that we have a set of *K* models which are characterized by having different subsets of z_t as predictors. Denoting these subsets by $z^{(k)}$ for *k = 1,..,K,* our set of models can be written as:

$$y_t = z_t^{(k)} \theta_t^{(k)} + \varepsilon_t^{(k)}$$
$$\theta_{t+1}^{(k)} = \theta_t^{(k)} + \eta_t^{(k)}$$

(1)

where y_t is the dependent variable being forecast, i.e., the return of the gold price, $\varepsilon_t^{(k)}$ is

$N(0, H_t^{(k)})$ and $\eta_t^{(k)}$ is $N(0, Q_t^{(k)})$. Let $L_t \in \{1, 2, .., K\}$ denote which model applies at each time

period, the coefficient vector be $\Theta_t = (\theta_t^{(1)'}, ..., \theta_t^{(K)'})'$ and the output vector be $y^t = (y_1, ..., y_t)'$.

The fact that we are allowing different models to hold at each point in time and performing

model averaging, gives rise to the terminology "dynamic model averaging". To be precise, when

forecasting time t variables using information through time t-1, DMA involves calculating the

probabilities $Pr(L_t = k \mid y^{t-1})$ for the models $k = 1, .., K$, and averaging forecasts across those K

models, using these probabilities. DMS involves selecting the single model with the highest value

for $Pr(L_t = k \mid y^{t-1})$ and using this model to forecast. However, there are problems with such a

framework, since many of the models can have a large number of parameters, and the

computational burden which arises when K is large implies that estimation can take a long time.

Thus, a full Bayesian approach to DMA can be quite difficult.

Following Koop and Korobilis (2012), in this paper we use approximations as suggested

by Raftery *et al.,* (2010). These approximations involve two parameters for the coefficients and

the models, λ and α, which Raftery *et al.,* (2010) refer to as the *forgetting factors,* and fix them to

numbers slightly below one. To explain the role of these forgetting factors, first consider the

standard state space model below for $t = 1....T$:

$$y_t = z_t \theta_t + \varepsilon_t$$

(2)

$$\theta_t = \theta_{t-1} + \eta_t$$

(3)

In our case, the output vector y_t is the return of the gold price, the vector

$z_t = [1, x_{t-1}, y_{t-1}, ..., y_{t-p}]$ is an $1 \times m$ vector of predictors for the return of the gold price which

also includes an intercept and lags of the dependent variable, the vector $\theta_t = [f_{t-1}, \beta_{t-1}, \gamma_{t-1}, ..., \gamma_{t-p}]$ is an $m \times 1$ vector of coefficients (states), and $\varepsilon_t \sim N(0, H_t)$ and $\eta_t \sim N(0, Q_t)$ are the errors which are assumed to be mutually independent at all leads and lags. For given values of the variance-covariance matrices H_t and Q_t, the standard filtering and smoothing results can be used to carry out recursive estimation or forecasting. Kalman filtering begins with the result that

$$\theta_{t-1} \mid y^{t-1} : N\left(\hat{\theta}_{t-1}, \Sigma_{t-1|t-1}\right) \tag{4}$$

where formulae for $\hat{\theta}_{t-1}$ and $\Sigma_{t-1|t-1}$ which are defined below are standard. Note here only that these formulae depend one variance-covariance matrices H_t and Q_t. for the errors. Then Kalman filtering proceeds, using:

$$\theta_t \mid y^{t-1} \sim N\left(\hat{\theta}_{t-1}, \Sigma_{t|t-1}\right), \tag{5}$$

where

$$\Sigma_{t|t-1} = \Sigma_{t-1|t-1} + Q_t.$$

Raftery *et al.* (2010) note that things get simplified substantially when one involves a forgetting factor (λ) in the state equation for the parameters. This is done by replacing this second equation of (5) by:

$$\Sigma_{t|t-1} = \frac{1}{\lambda} \Sigma_{t-1|t-1} \tag{6}$$

or, equivalently, $Q_t = (1 - \lambda^{-1})\Sigma_{t-1|t-1}$ where $0 < \lambda \leq 1$. The term "forgetting factor" is suggested by the fact that this specification implies that observations j periods in the past have weight λ^j. It can also be noted that it implies an effective window of $1/(1-\lambda)$. In the literature, it is common to choose a value of λ near one, which suggests a gradual evolution of coefficients. Raftery *et al.* (2010) set $\lambda = 0.99$. For monthly macroeconomic data, this suggests observations

five (one) years ago receive approximately 50% (90%) as much weight as last period's observation. This would be consistent with fairly stable models where the coefficient change is gradual. This would suggest substantial parameter instability with rapid change in coefficients. Forecasting in the one model case is then completed by the updating equation:

$$\theta_t \mid y^t \sim N(\hat{\theta}_t, \Sigma_{t|t}),\tag{7}$$

where

$$\hat{\theta}_t = \hat{\theta}_{t-1} + \Sigma_{t|t-1} z_t \left(H_t + z_t \Sigma_{t|t-1} z_t' \right)^{-1} \left(y_t - z_t \hat{\theta}_{t-1} \right)\tag{8}$$

and

$$\Sigma_{t|t} = \Sigma_{t|t-1} - \Sigma_{t|t-1} z_t \left(H_t + z_t \Sigma_{t|t-1} z_t' \right)^{-1} z_t \Sigma_{t|t-1}.\tag{9}$$

Recursive forecasting is done using the predictive distribution

$$y_t \mid y_{t-1} \sim N\left(z_t \hat{\theta}_{t-1}, H_t + z_t \Sigma_{t|t-1} z_t' \right).\tag{10}$$

For the case with many models, (1), uses the previous approximation and an additional one. To understand this, we now switch to the notation for the multiple model case in (1) and let Θ_t denote the vector of all the coefficients. In the standard single model case, Kalman filtering is based on (4), (5) and (7). In the multi-model case, for model k, these three equations become:

$$\Theta_{t-1|L_{t-1}} = k, y^{t-1} \sim N\left(\hat{\theta}_{t-1}^{(k)}, \Sigma_{t-1|t-1}^{(k)} \right)\tag{11}$$

$$\Theta_{t|L_t} = k, y^{t-1} \sim N\left(\hat{\theta}_{t-1}^{(k)}, \Sigma_{t|t-1}^{(k)} \right)\tag{12}$$

$$\Theta_{t|L_t} = k, y^t \sim N\left(\hat{\theta}_t^{(k)}, \Sigma_{t|t}^{(k)} \right),\tag{13}$$

where $\hat{\theta}_t^{(k)}$, $\Sigma_{t|t}^{(k)}$ and $\Sigma_{t|t-1}^{(k)}$ are obtained via Kalman filtering in the usual way using (6), (8), and (9), except with (k) superscripts added to denote model k.

The previous results are all conditional on $L_t = k$, and we need a method for unconditional prediction. In this paper, we follow the suggestion of Raftery *et al.* (2010) and as

used in Koop and Korobilis (2012), based on a forgetting factor α for the state equation for the models which in turn, is, comparable to the forgetting factor λ used with the state equation for the parameters. The derivation of Kalman filtering ideas begins with (4). The analogous result, when doing DMA, is

$$P\left(\Theta_{t-1}, L_{t-1} \mid y^{t-1}\right) = \sum_{k=1}^{K} P\left(\theta_{t-1}^{(k)} \mid L_{t-1} = k, y^{t-1}\right) Pr\left(L_{t-1} = k \mid y^{t-1}\right), \tag{14}$$

where $P(\theta_{t-1}^{(k)} \mid L_{t-1} = k, y^{t-1})$ is given by (11). To simplify notation, let $\pi_{t|s,l} = Pr(L_t = l \mid y^s)$ and thus, the final term on the right hand side of (14) is $\pi_{t-1|t-1,k}$.

With,

$$\pi_{t|t-1,k} = \frac{\pi_{t-1|t-1,k}^{\alpha}}{\sum_{l=1}^{K} \pi_{t-1|t-1,l}^{\alpha}}, \tag{15}$$

where $0 < \alpha \leq 1$ is set to a fixed value slightly less than one and is interpreted in a similar manner to λ, i.e., if $a = 0.99$ (our benchmark value and also the value used by Raftery *et al.*, 2010) the forecast performance five years ago receives 50% as much weight as the forecast performance of last period (when using monthly data), while the forecast performance one year ago receives about 90% as much weight as last month's performance. Comparable to those of the updating equation in the Kalman filter, we have a model updating equation of:

$$\pi_{t|t,k} = \frac{\pi_{t|t-1,k} Pk(y_t \mid y^{t-1})}{\sum_{l=1}^{K} \pi_{t|t-1,l} Pl(y_t \mid y^{t-1})}, \tag{16}$$

where $p_l(y_t \mid y^{t-1})$ is the predictive density for model l, which is the Normal density in (10) with (l) superscripts added and evaluated at y_t. Recursive forecasting can be done by averaging over predictive results for every model using $\pi_{t|t-1,k}$. So, for instance, DMA point predictions are given by:

$$E(y_t \mid y^{t-1}) = \sum_{k=1}^{K} \pi_{t|t-1,k} z_t^{(k)} \hat{\theta}_{t-1}^{(k)}.$$

DMS proceeds by selecting the single model with the highest value for $\pi_{t|t-1,k}$ at each point in time and simply using it for forecasting. Note also that, if $a = 1$, then $\pi_{t|t-1,k}$ is simply proportional to the marginal likelihood using data through time t-1, and yields the standard approaches to BMA. If we also set $\lambda = 1$, then we obtain BMA using conventional linear forecasting models with no time variations in coefficients. In our forecast comparison exercise, we include BMA in our set of alternative forecasting procedures and implement this by setting $a = \lambda = 1$.

The preceding discussion is all conditional on H_t. Raftery *et al.* (2010) recommend a simple plug in method where $H_t^{(k)} = H^{(k)}$ and is replaced with a consistent estimate. When forecasting the gold return, however, it is likely that the error variance is changing over time. Theoretically, we could use a stochastic volatility or an ARCH specification for $H_t^{(k)}$. However, this is computationally burdensome, so instead we follow a simple plug-in approach as in Koop and Korobilis (2011), which in turn, is a rolling version of the recursive method of Raftery *et al.* (2010). Specifically, let

$$\tilde{H}_t^{(k)} = \frac{1}{t^*} \sum_{j=t-t^*+1}^{t} \left[\left(y_t - z_t^{(k)} \hat{\theta}_{t-1}^{(k)} \right)^2 - z_t^{(k)} \Sigma_{t|t-1}^{(k)} z_t^{(k)\prime} \right] \tag{17}$$

To allow for more substantial changes in the error variances, we set $t^* = 20$ and, thus, use a rolling estimator based on five years data. Following Raftery *et al.* (2010), we can avoid the rare possibility that $\tilde{H}_t^{(k)} < 0$, by replacing $H_t^{(k)}$ by $\hat{H}_t^{(k)}$ where

$$\hat{H}_t^{(k)} = \begin{cases} \tilde{H}_t^{(k)} & \text{if} \quad \tilde{H}_t^{(k)} > 0 \\ \hat{H}_{t-1}^{(k)} & \text{otherwise} \end{cases}.$$

4. Data description

As indicated earlier, the data include six groups of predictors of the gold return. These predictors are categorized as the business cycle factor: the nominal factor; the interest rate factor; the commodity factor, the exchange rate market; and the stock price factor. In addition to these factors, we added the U.S. economic uncertainty and the Kansas City Fed's financial stress index to capture the impacts of policy uncertainty and anxiety in the economy. As indicated earlier, each factor is created based on the first principal component of the variables in that group. The business cycle factor, for example, is represented by the industrial production for the United States, the European Union (EU), China and Japan. An increase in this factor implies increases in the demand for jewelry which makes up more than 65% of total gold demand. The same description applies to the other factor.

One of the biggest advantages with the use of principal component in our study is that it greatly reduces number of parameters to estimate. This issue becomes crucial when we allow those parameters to vary over time. Based on the data availability, especially the availability of the China data, the sample period is January 1992 – November 2013 (263 monthly observations) and the detailed list and the sources of the variables are available in Table A1 in the Data Appendix. Note that, due to data transformation to ensure stationarity, which is required for both extracting factors and convergence issues in time-varying models, our effective sample starts from February, 1992, i.e., 1992:02, and goes until 2013:11.

The principal component analysis reveals that the first principal component in each group explains: 39 percent of the total variation in the business cycle group, 30 percent in the nominal group, 61 percent in the interest rate group, 69 percent in the commodity price group, 74 percent in the exchange rate group, and 73 percent in the stock price group. Figure A1 in the

Data Appendix plots the normalized growth rates of variables and their common factors extracted based on the their economic relevance. [3]

Kansas City Federal Reserve Bank's financial stress index (KC*FSI*), which is not part of the first principal components, measures the degree of financial stress in the financial markets and is constructed from 18 data financial series, where each of these series captures some aspect of financial stress. This index is accessed from the Federal Reserve Bank of Saint Louis's database. The U.S. economic policy uncertainty index (EPU)[4] is introduced by Baker *et al.* (2012). It is constructed from three types of underlying components. The first component quantifies newspaper coverage of policy-related economic uncertainty. The second reflects the number of federal tax code provisions set to expire in future years. The third uses disagreement among economic forecasters as a proxy for uncertainty. A weighted normalized index is created from those three components. This index spiked on October 3, 2008 which is the depth of the GFC.**5.**

5. Empirical Results

We present the empirical results in two sub-sections, using DMA and DMS as in Koop and Korobilis (2011), i.e., by setting $\alpha = 0.99$ and $\lambda = 0.99$ [5], a non-informative prior over the models ($\pi_{0|0,k} = 1/K$ for $k = 1, ..., K$) and a relatively diffuse prior for the initial state conditions: $\theta_0^{(k)} \sim N(0,100)$ for $k = 1, ..., K$. The first sub-section provides evidence of which variables are good predictors of the gold price returns from the list of the eight potential factors and variables as categorized and listed in the Data Appendix. In the second sub-section, we compare the

---

[3] One could include more than one principal component in each group. For example, adding a second principal component in the business cycle category will improve the explanatory power by 24 percentage points. However, it will be difficult to interpret or to give economic meanings. To give better economic meaning to the factors, we therefore extract one common factor from each group.

[4] See http://www.policyuncertainty.com

[5] Note that, one could choose values for λ and α based on the forecast performance as in Grassi and Santucci de Magistris (2013), but this would bias our results in favour of DMA and thus is not a valid procedure for out-of-sample forecasting (Koop and Korobilis, 2011). Alternatively, when forecasting at time τ, we could consider a grid of values for λ and α, and select the value which yielded the highest value for the marginal likelihood or an information criterion, which essentially amounts to treating λ and α as unknown parameters. However, this would greatly add to the computational burden; so much so that it might be impossible to do forecasting in real time (Koop and Korobilis, 2011). Hence, we follow Koop and Korobilis (2011, 2012) and simply select values for the forgetting factors.

forecast performance of the DMA to those of a number of alternative forecasting models nested in the DMA, including the BMA, TVP and random walk models. The considered forecast horizons are: one ($h=1$), three $(h=3)$, six ($h=6$), nine (h=9) and 12 ($h=12$) months. All of the models include an intercept and a zero lag of the dependent variable as chosen by the Schwarz information criterion.

5.1. *Good predictors for the gold price returns*

One of the largest potential advantages of using the DMA framework is that it allows the forecasting model to change over time, i.e. the model parameters may change as well as the set of predictors. Given the difficulty in the literature in explaining the gold price return movements, we include a set of eight possible predictors (excluding the intercept), hence we have 256 possible models to choose from.

Turning to the results on the main predictors as selected by the DMA analysis for the gold returns for every forecast horizon, $h=1$, $h=3$, $h=6$, $h=9$ and $h=12$, over the out-of-sample period 1999:10 to 2013:11. The posterior probabilities of inclusion of each predictor for the different forecast horizons, which shed light on their predictive abilities, are presented in Figures 2 and 3. It appears that the good predictors of gold price returns vary over time and over the forecast horizons. Particularly, three periods are striking namely pre-2005, 2005-2009 and post 2009. It must be noted that the middle period includes the heart of the boom for commodities and the Great Recession period, while the third is the post Great Recession or recovery period.

Prior to 2005, the exchange rate factor (*EXCH*) comes through strongly for all horizons except for $h=9$, while the interest rate (*INTEREST*) shows strong predictive power throughout all forecast horizons. The business cycle factor (*BUSCYCLE*) comes through strongly for all horizons but $h =3$. We observe US economic policy uncertainty (*EPU_US*) playing a strong predictive role for $h= 3$ and $h=12$ and nominal factors doing the same for $h=1$ and $h=3$. The nominal factor (*NORMFAC*) and the stock market factor (*STOCK*) possess strong predictive

power for $h = 1$ and 2, and $h=1$ and 12, respectively. The commodity prices factor (*COMMPRICE*) comes through strongly for $h=1$ only. The Kansas City Fed' Financial Stress Index (*KCFSI*) shows strong predictive power for all horizons except for h=6. In sum, in the pre-2005 period the exchange rate, the financial stress index, interest rate and business cycle factors are the leaders of the forecasting procession.

Between 2005 and 2009, the same pattern is observed especially with *EXCH* and *INTEREST* having strong predictive power for gold at all horizons except for $h=9$ and *STOCK* at $h=1$ and 12. A similar pattern is also observed for *BUSCYCLE* as it proves strong for $h= 1$, 6 and 9 and *COMMPRICE* for $h=1$. However, *KCFSI* emerges as an important predictor for only $h=3$, 9 and 12, while *EPU_US* comes through strongly for $h=3$ and 9. This period reinforces the results of the previous period which bestow good predictive power on the exchange rate and interest rate and the financial stress to a lesser extent. Gold is known as a currency with high monetary value and the interest rate is the opportunity cost of holding gold.

For the post 2009 period, we observe that *EXCH* and *STOCK* maintain the same predictive roles as in the previous sub-period. *INTEREST* comes through strongly for $h=6$ and 12, while *EPU_US* comes through strongly for $h=3$ and 9 as in the 2005-2009 sub-period. *COMMPRICE* and *KCFSI* have predictive power for only $h= 3$ and $h=12$, respectively. *BUSCYCLE* shows strong predictive power for $h=6$, 9 and 12. Overall, all the predictors show strong predictive power at one time or another though at varying magnitudes, while *EXCH* and *KCFSI* appear to be strong at almost all horizons and sub-periods. However, the forecasting prowess of the exchange rate is supreme.

5. 2. *Model forecasting comparison*

To analyze the forecast performance of each model, we use the mean squared forecast error (MSFE) in percentages, which is available for all the models. Note that, the preferred method for the Bayesian forecast comparisons is the sum of the log predictive likelihoods, involving the

entire predictive distribution. For a discussion on predictive likelihoods, refer to Gweke and Amisano (2007) as a representative source. The predictive likelihood is the predictive density for y_t, given data through time *t-1*, evaluated at the actual outcome. Since we use the direct method of forecasting, the log predictive density for the *h*-step ahead forecast is an obvious extension of the formula for the one-step ahead predictive density in model *l* as denoted by $p_l(y_t | y^{t-1})$ and described in Section 3. We use the sum of log predictive likelihoods for the forecast evaluation of the Bayesian models, with the sum beginning in the tenth month of 1999 up to the eleventh month of 2013 of the out-of-sample period. MSFE's are reported over the same period.

We present results for the following alternative forecasting models:

(1) Forecasts using Dynamic Model Averaging (DMA) with $\alpha = \lambda = 0.99$.

(2) Forecasts using Dynamic Model Selection (DMS) with $\alpha = \lambda = 0.99$.

(3) Forecasts using all the variables in a single model with time varying parameters (TVP) (100% of the prior weight is attached to the model with all the variables in this special case of DMA, with all other modeling choices being identical including λ=0.99).

(4) Forecasts using DMA, but the coefficients do not vary over time (a special case of DMA where $\lambda = 1$ and $\alpha = 0.99$).

(5) Forecasts using Bayesian Model Averaging (BMA: a special case of DMA where $\alpha = \lambda = 1$).

(6) Forecasts using the random walk (RW).

Model (1) applies the dynamic model averaging (DMA), using the benchmark values $\alpha = \lambda = 0.99$. This allows not only for the parameters to change over time, but also for the set of predictors. As stated earlier, values of 0.99 are consistent with fairly stable models with gradual coefficient changing over time, where forecasts are averages across models using the associated probabilities calculated as $Pr(L_t = k | y^{t-1})$. Model (2), where dynamic model selection (DMS) is used, involves selecting the single model with the highest probability and using this to

forecast. Model (3) is a special case of Model (1) where all predictors are included for all time periods, but the parameter values are allowed to change. Model (4) is a constant parameter model but the model evolution is allowed, while in Model (5), through setting both forgetting factors to one, we obtain the Bayesian model averaging (BMA) which uses conventional linear forecasting models with no time variations in the coefficients. Understandably, the random walk (RW) model is non-Bayesian, which means that there is no predictive likelihood for the model.

Table 1 presents the results from our forecasting exercise for the gold price return for five forecasting horizons, namely 1-month-ahead (Horizon 1) , 3-months-ahead (Horizon 3), 6-months-ahead (Horizon 6), 9-months-ahead (Horizon 9), and 12-months-ahead (Horizon 12). From the results in Table 2, it is clear that DMA and DMS forecast generally well, with DMS being the best overall. The sum of the log predictive likelihoods clearly shows that DMS performs better across all forecast horizons. This result also carries over to the MSFE. Although DMS and DMA can be interpreted as doing the shrinkage in different ways, DMS puts weight on all models other than the one best model, thus shrinking the contribution of all models except one towards zero. This additional shrinkage appears to have given DMS additional benefits over DMA. This finding is consistent with forecasts obtained by Koop and Korobilis (2012) for inflation and Gupta et al. (forthcoming) for China's foreign exchange reserves using similar models. Both the sum of the log predictive likelihoods as well as the MSFE indicates that using a TVP model for forecasting results in poorer forecasting performance, relative to the DMA and DMS. The poor forecast performance of the TVP model compared to DMA and DMS is an indication that the shrinkage provided by the latter models is of great value in forecasting. Also note that, the predictors matter in forecasting, since all the models that include the predictors outperforms the RW model. However, to obtain best forecast performances, one not only needs to include information from the fundamentals, but also allows for model evolution and parameter evolution as carried out by the DMA and DMS. Interestingly though, based on the results provided by the predictive likelihoods, we see that most of the

improvements in forecast performance found by DMA or DMS are due to model evolution rather than parameter evolution since DMA and DMS perform better than the DMA model with λ set equal to 1 (implying the model is not allowed to change here). It can therefore be concluded that allowing for model uncertainty and not only parameter uncertainty, improves the forecasting performance of these models.

6. Conclusion

Given the difficulty in the literature in forecasting the gold price returns, this study provides evidence regarding which variables are good predictors of those returns from the list of the eight potential factors and variables. Then it compares the forecast performance of the DMA model to those of a number of alternative forecasting models nested in DMA, including the BMA and the TVP model as well as the random walk model. One of the important advantages of DMA is that it allows the model parameters and the set of predictors to change over time. The forecast horizons are one, three, six, nine and 12 months. The performance of the good predictors of the gold returns varies over time and over the forecast horizons, highlighting the importance of differences in three periods. The first period ranges from 1992 to 2005, which includes two major wars, steady increases in gold prices and a part of the commodity boom. The second period covers the years 2005-2009, which contains the last part of the commodity boom and the Great Recession. The third is the post 2009 period which comprises the recovery and relative declines in gold prices.

The results attest to the strong predictive power of the exchange rates in forecasting the gold returns over most forecast horizons under the three time periods. Gold is considered a resource currency, and a hedge and a safe against inflation and crises during which national currencies including the dollar lose value. Two other financial variables with good predictive power include the interest rate factor and the financial stress index (FSI). These variables play a good role in predicting the gold price returns particularly during the first two time periods, the pre-2005 and

the 2005-2009 periods. The economic activity factor has a good predictive power in the first two time periods but not in the third period which is the recovery period. This may reflect the low economic growth which might not make economic activity a strong predictor in this period. There is also clear evidence that the stock market has a good predictive power in the third period, in addition to that possessed by the exchange rate factor. This may have to do with the strong responses of different asset classes to increased liquidity as a result of the Fed's quantitative easing that took place in this period.

Moreover, to obtain the best forecast performance, one not only needs to include information from the fundamentals, but also allows for model evolution and parameter evolution as supported by the DMA and DMS. However, the empirical evidence supports DMS over DMA. In sum, allowing for model uncertainty and not only parameter uncertainty, improves the forecasting performance of the models.

The major conclusion of this paper is that in the post Great Recession or the slow economic recovery period, the variables or factors that have the strongest forecasting power are the exchange rates and the stock markets. However, the economic activity factor lost some of its forecasting glitter over this period as it does not have a strong predictive power as the exchange rates and the stock markets . This interesting result underscores the importance of increasing financialization of commodities including gold in comparsion to the business cycle variables, again probably as a result of slow economic growth and prolonged quantitative easing. This conclusion became evident from the important performance of models that takes in consideration parameter and model uncertainties.

Table 1: Comparing different forecasting methods for gold price returns

| Forecast Method	Sum of Log predictive likelihoods	MSFE
h=1		
DMA	-471.125	19.635
DMS	-462.856	16.918
TVP (α= 1)	-474.7	20.287
DMA (λ=1)	-470.821	19.646
BMA (DMA with α=λ=1)	-473.990	20.221
Random walk	-	31.253
h=3		
DMA	-499.201	25.113
DMS	-486.544	22.467
TVP (α= 1)	-508.663	27.004
DMA (λ=1)	-502.994	26.140
BMA (DMA with α=λ=1)	-511.597	28.194
Random walk	-	29.839
h=6		
DMA	-475.304	23.940
DMS	-463.958	18.996
TVP (α= 1)	-477.788	24.381
DMA (λ=1)	-477.580	24.373
BMA (DMA with α=λ=1)	-483.630	25.620
Random walk	-	33.123
h=9		
DMA	-458.793	18.937
DMS	-454.973	17.788
TVP (α= 1)	-465.219	19.497
DMA (λ=1)	-460.791	18.994
BMA (DMA with α=λ=1)	-464.933	19.531
Random walk	-	33.480
h=12		
DMA	-437.909	18.195
DMS	-423.748	16.172
TVP (α= 1)	-445.222	19.310
DMA (λ=1)	-438.273	18.799
BMA (DMA with α=λ=1)	-442.483	19.291
Random walk	-	33.476
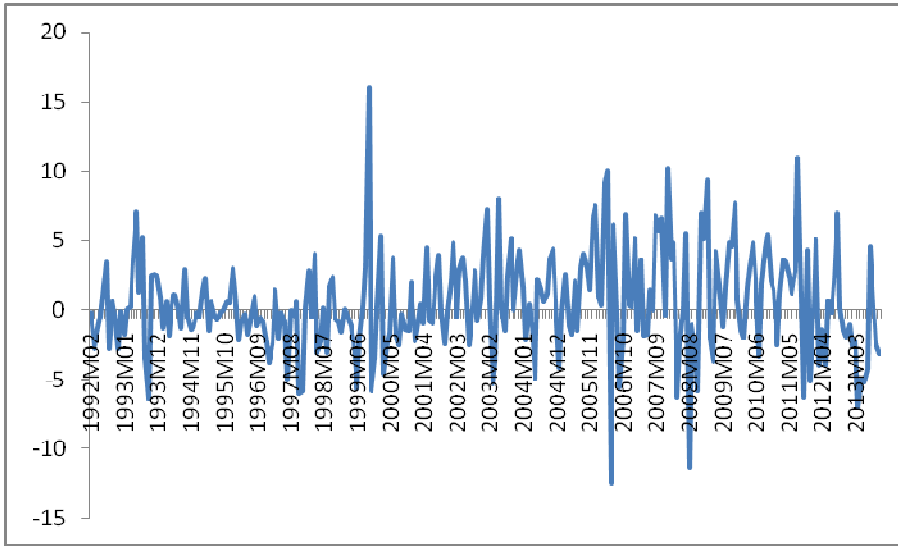
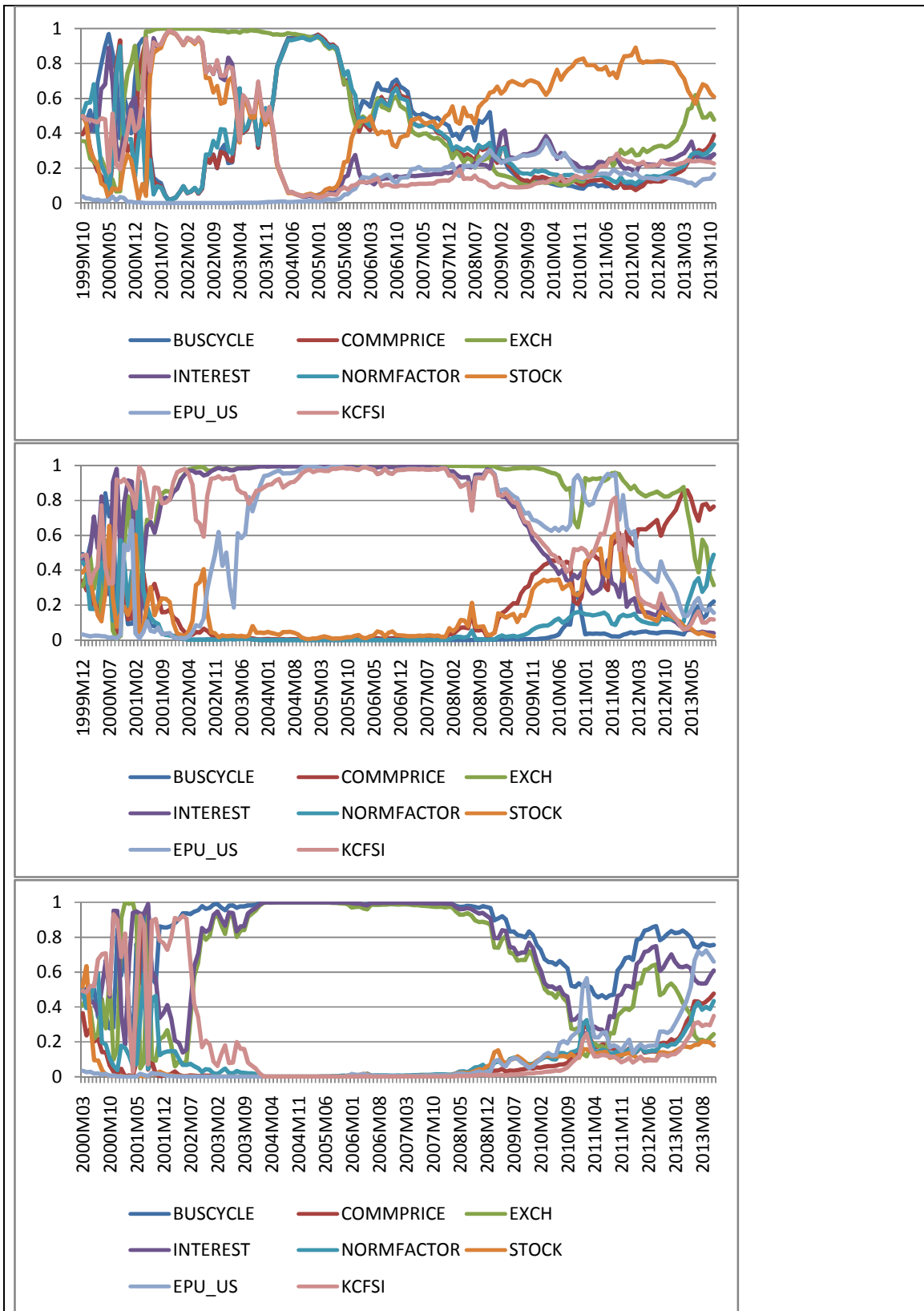Figure 1: Returns on gold (in percentage, month-on-month)

Figure 2: Posterior inclusion probabilities of predictors (upper panel: $h=1$; middle panel: $h=3$; lower panel: $h=6$)
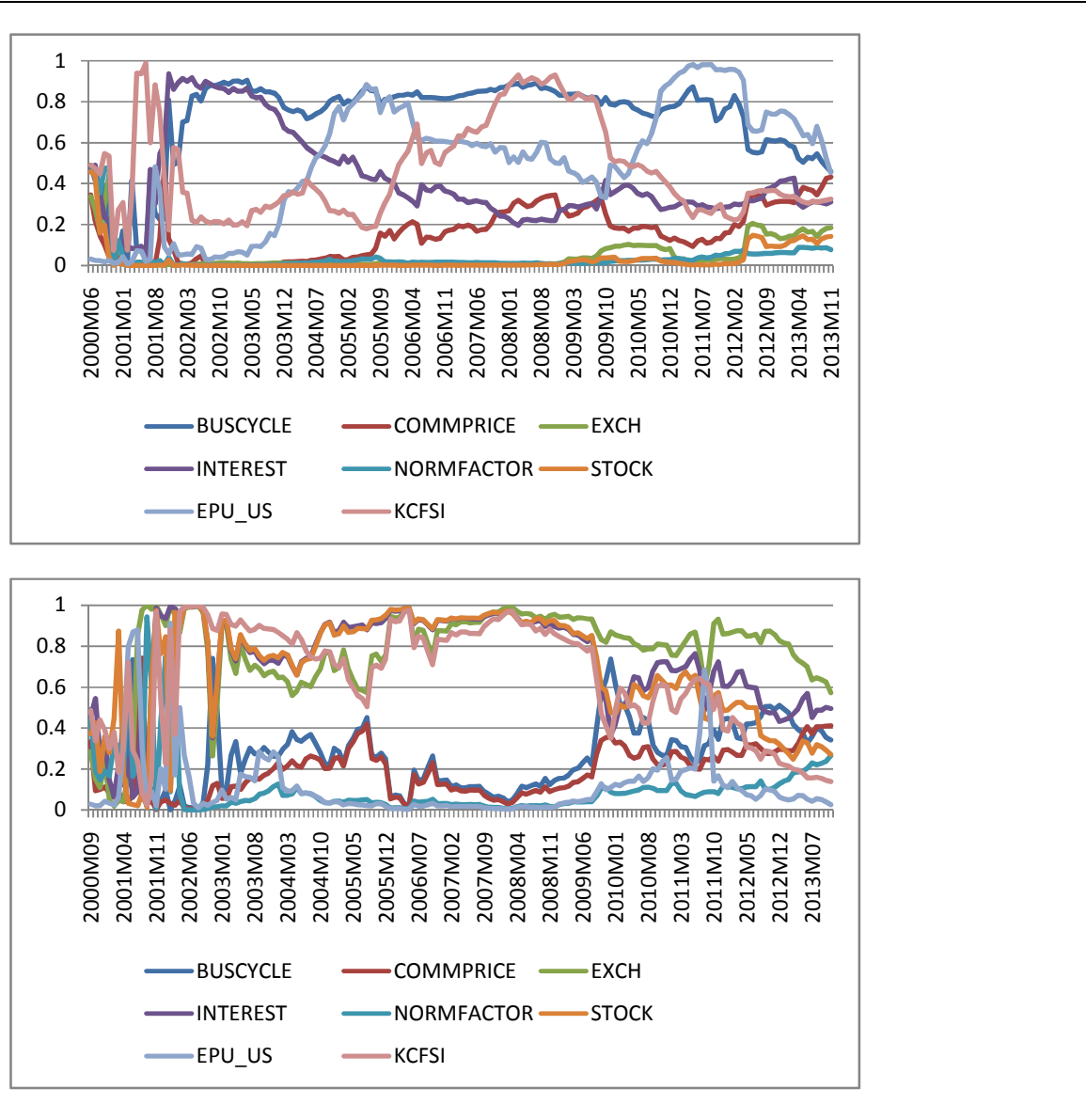
Figure 3: Posterior inclusion probabilities of predictors (upper panel: h=9; lower panel: h=12).

References

Apergis, N. (2014). Can gold prices forecast the Australian dollar movements*? International Review of Economics & Finance* 29, 75-82.

Baur, D. G. and Glover, K. J. (2014). Heterogeneous expectations in the gold market: Specification and estimation. *Journal of Economic Dynamics and Control* 40, 116-133.

Ciner, C., Gurdgiev, C., Lucey, B. M. (2013). Hedges and safe havens: An examination of stocks, bonds, gold, oil and exchange rates. *International Review of Financial Analysis* 29, 202–211.

Geweke, J. and G. Amisano (2011). Hierarchical Markov Normal Mixture Models with Applications to Financial Asset Returns. *Journal of Applied Econometrics* 26, 1-29.

Grassi, S. and Santucci De Magistris, P. (2013). It's all about volatility (of volatility): evidence from a two-factor stochastic volatility model. *CREATES Research Paper* 2013-03.

Gupta, R.., Hammoudeh, S., Kim, W.J. and Simo-Kengne, B.D. (2014) Forecasting China's foreign exchange reserves using dynamic model averaging: The role of macroeconomic fundamentals, financial stress and economic uncertainty. *North American Journal of Economics and Finance (in press).*

Ismail, Z., A., Yahya, A. and Shabri, A. (2009). Forecasting gold prices using multiple linear Regression method. *American Journal of Applied Sciences* 6 (8): 1509-1514, 2009.

Koop, G. and D. Korobilis (2011). UK macroeconomic forecasting with many predictors: Which models forecast best and when do they do so? *Economic Modelling* 28, 2307-2318.

Koop, G. and D. Korobilis (2012). Forecasting inflation using dynamic model averaging. *International Economic Review* 53, 867-886.

Lili, L. and Chengmei, D. (2013). Research of the influence of macroeconomic factors on the price of gold. *Procedia Computer Science* 17, 737-743.

Pindyck, R. S. and Rotemberg, J. J. (1990). The excess co-movement of commodity prices. *The Economic Journal* 100, 1173-1189.

Pierdzioch, C., Risse, M. and Rohloff, S. (2014a). On the efficiency of the gold market: Results of a real-time forecasting approach. *International Review of Financial Analysis* 32, 95–108.

Pierdzioch, C., Risse, M. and Rohloff, S. (2014b). The International business cycle and gold-price fluctuations. *The Quarterly Review of Economics and Finance* (in press).

Raftery, A.E., Kárný, M. and P. Ettler (2010). Online prediction under model uncertainty via dynamic model averaging: application to a cold rolling mill. *Technometrics* 52(1), 52-66.

Sjaastad. L. J. (2008). The price of gold and the exchange rates: Once again. *Resources Policy.* 33, 118–124.

Sjaastad, L. J and Scacciavillani, F. (1997). The price of gold and the exchange rate. *Journal of International Money and Finance* 15 (6), 879-897.

Data Appendix

Table A1: Variables used in predicting the gold price returns

| Category	Tcode*	Description	Source
Business Cycle Factor (BUSCYCLE)	4	EU Industrial Production	Global Insight
	4	China Industrial Production	China National Bureau of Statistics
	4	US Industrial Production	FRED
	4	Japan Industrial Production	Ministry Of Economy, Trade And Industry
Nominal Factor (NORMFACTOR)	4	EU CPI	Global Insight
	4	China CPI	China National Bureau of Statistics
	4	US CPI	Bureau of Labor Statistics
	4	Japan CPI	Statistics Bureau
	4	EU M1 Index	Global Insight
	4	US M1 (Bill. USD)	FRED
Interest Rate Factor (INTEREST)	4	EU Policy Rate	Global Insight
	4	1 Month LIBOR on British Pounds	British Bankers' Association
	4	US Policy Rate	FRED
Commodity Price Factor (COMMPRICE)	4	CRB US Spot Commodity Price Index	Datastream
	4	All Commodity Price Index	IMF
	4	Dated Brent, light blend 38 API, fob U.K.	IMF
	4	Dubai, medium, Fateh 32 API, fob Dubai Crude Oil	IMF
	4	West Texas Intermediate 40 API, Midland Texas	IMF
	4	Silver Fixing Price (US cents per troy ounce)	London Bullion Market Association
Exchange Rate Factor (EXCH)	4	US effective exchange rate (foreign/dollar)	FRED
	4	Yen/dollar	FRED
	4	dollar/euro	Global Insight
Stock Market Factor (STOCK)	4	The STOXX Europe 600 Index represents large, mid and small capitalization companies across 18 countries of the European region.	STOXX (www.stoxx.com)
	4	The STOXX North America 600 Index represents the largest companies in the Americas region.	STOXX (www.stoxx.com)
	4	The STOXX Asia/Pacific 600 Index represents the largest companies in Australia, Hong Kong, Japan, New Zealand and Singapore.	STOXX (www.stoxx.com)
	4	MSCI_emerging	MSCI (www.msci.com)
KCFSI	1	The Kansas City Fed's Financial Stress Index	Kansas City Fed
EPU_US	1	Economic Policy Uncertainty Index for United States	Scott Baker, Nicholas Bloom and Steven J. Davis

Note: 1 and 4 in Tcode (transformation code) indicate no transformation (level) and the first-difference of the logarithm of the variable, respectively.

Figure A1: Standardized growth rates and the common factors
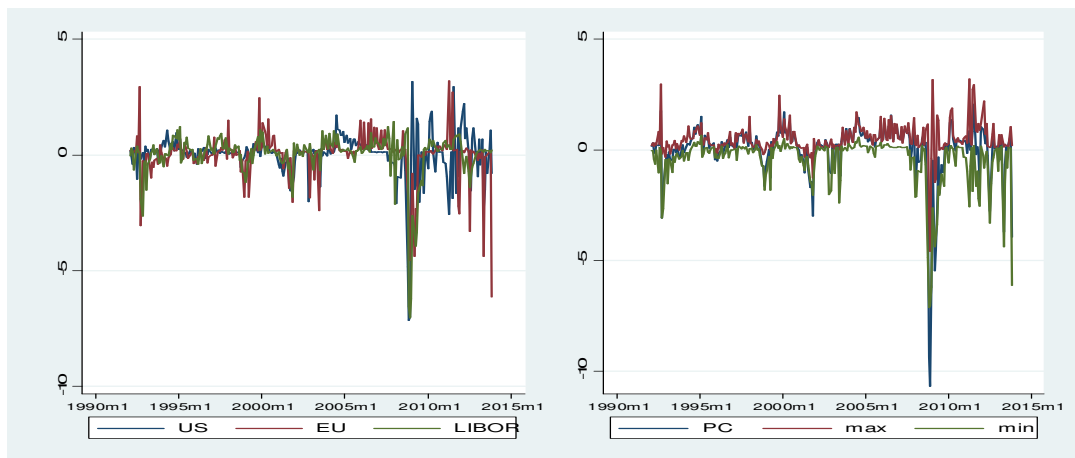
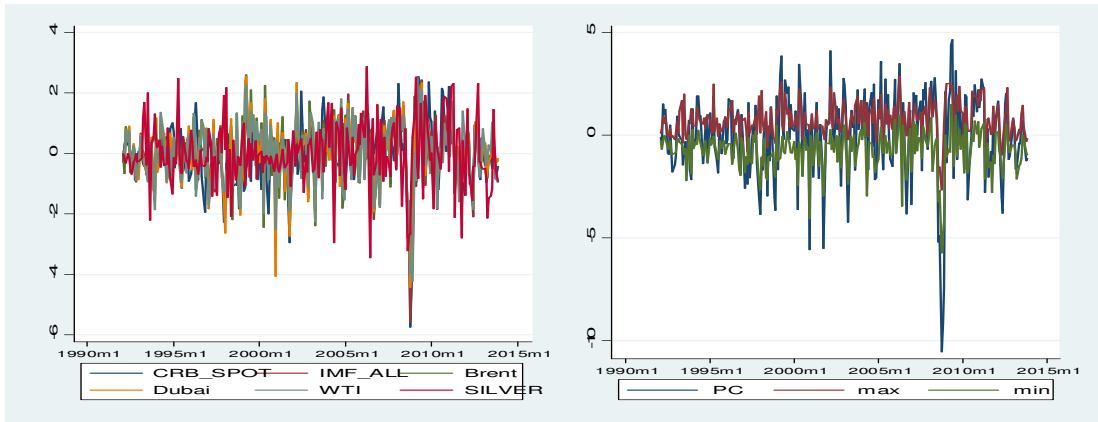A. Business Cycle Factor*
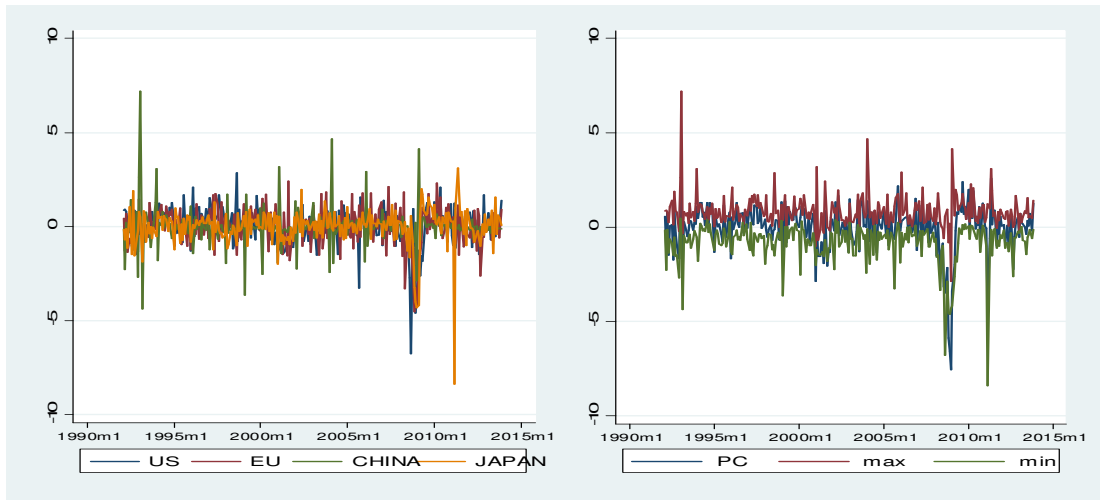
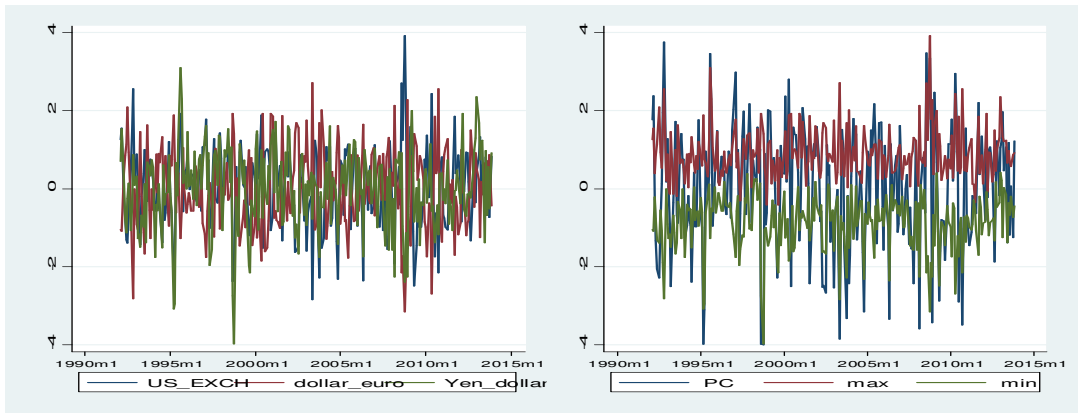B. Nominal Factor
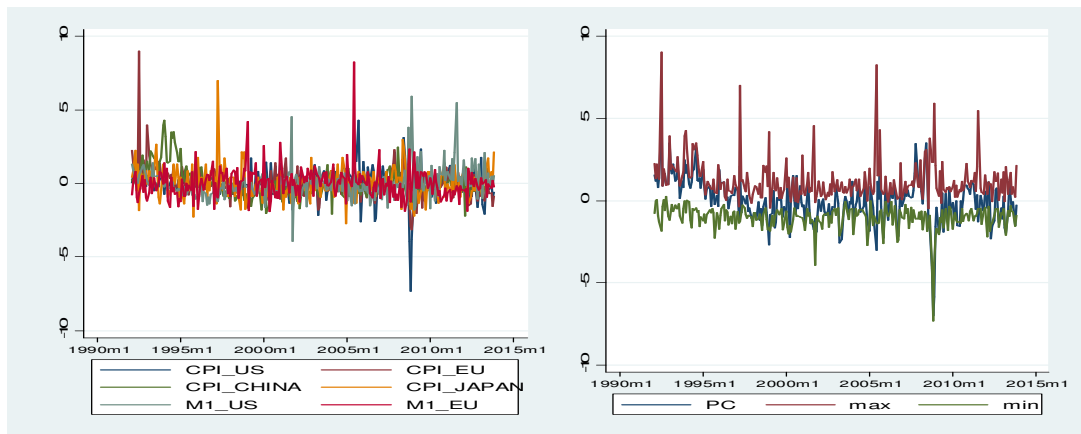
C. Interest Rate Factor

* PC denotes the principal component, max (min) denote the max (min) of the growth rates within the group. Growth rate is expressed in percent (%).

Figure A1: Standardized growth rates and the common factors (continued)

D. Commodity Price Factor

E. Exchange Rate Factor

F. Stock Price Factor