Exchange Rate Forecasting with Structural Shocks as Predictors *

Konstantin Styrin

This draft: November 28, 2008

JOB MARKET PAPER

Abstract

As documented by many studies, monetary policy (MP) shocks account for a considerable fraction in forecast error variance decomposition of exchange rates. Impulse responses of exchange rates to identified MP shock are characterized by “delayed overshooting” with peak effect in 1 to 3 years. This implies that estimated MP shock should have a non-trivial forecasting value with respect to exchange rates. I examine this conjecture by forecasting exchange rates of 9 non-Euro OECD currencies against the US dollar out of sample at horizons up to 24 months with the US MP shock as predictor. The MP shock is identified and estimated in a multi-country Factor Augmented VAR using a block-recursive identification scheme. I do not find any evidence that forecasts with the US MP shock tend to robustly outperform a driftless random walk at any horizons. However, partially identified group of shocks that contemporaneously affect mostly financial market variables are shown to be good predictors for exchange rates of commodity exporters 4-24 months ahead. I interpret these shocks as news about future prospects of the US economy.

JEL Classification: C53, F31, F37

Key words: Exchange rates; Forecasting; Factor-augmented vector autoregressions; Structural shocks.

1 Introduction

It is well known that structural macroeconomic models have had hard times in producing out-of-sample forecasts of exchange rates superior to a nave driftless random walk (RW) benchmark at short horizons (Meese and Rogoff, 1983). The majority of forecasting specifications in the literature exploit cointegration relationship between the exchange rate and its fundamentals as the main source of forecastability. The cointegration implies that, over time, the exchange rate converges to the value determined by fundamentals such as relative money supplies, interest rate differentials, etc. Existing models of this class are able to outperform RW robustly across currencies and time samples only at horizons longer than 2-4 years. Their performance at shorter horizons remains poor.

*I am grateful to Kenneth Rogoff and James Stock for invaluable advice and encouragement and Harvard International Seminar participants for helpful comments. All errors are mine.
This study takes a different approach to exchange rate forecasting and uses (partially) identified and estimated structural shocks as predictors. The shocks are identified and estimated within a multi-country Structural Factor Augmented Vector Autoregression (Structural FAVAR, or SFAVAR) model.

The main structural shock of interest is monetary policy (MP) shock in the US. A number of studies based on structural vector autoregressions (SVAR) document that impulse response of exchange rate to identified MP shock is hump-shaped (Eichenbaum and Evans, 1995; Faust and Rogers, 2003; Scholl and Uhlig, 2007). This phenomenon is called “delayed overshooting” effect. A surprise monetary contraction in the US leads to nominal appreciation of the US dollar against other currencies in the long run. In the short run, the US exchange rate overshoots its new long-run equilibrium level. According to the above cited studies, the peak appreciation occurs at horizons one to three years (Eichenbaum and Evans, 1995). After that the US currency starts depreciating and exchange rate gradually converges to its new long-run equilibrium level. The MP shock accounts for 20-40% of exchange rate forecast error variance decomposition.

“Delayed overshooting” implies that the MP shock should be able to forecast exchange rate at horizons around the peak impulse response if the following three conditions hold. First, the hump-shaped impulse response is not a consequence of misidentified MP shock. Second, the peak effect is strong enough quantitatively. Third, the MP shock can be estimated with a sufficient degree of precision.

Reliable identification of the MP shocks is critical for the purposes of this study. Misidentified MP shock will be useless in forecasting exchange rates out of sample and bring nothing but noise. FAVAR is believed to have a potential to better identify and estimate structural shocks compared with conventional small-size VAR. One of the main limitations of VAR is their vulnerability to “dimensionality curse”: the number of parameters to be estimated grows as the number of included variables squared. That is why, in practice, typical VAR’s have only a few variables. FAVAR circumvents this difficulty by assuming that all included variables are driven by a relatively small number of common shocks. This makes it capable to operate with large macro panel of data.

FAVAR thus sets the econometrician on an equal ground with economic agents and policy makers whose information sets are much larger than a standard VAR can accommodate. This should eliminate omitted variable bias and result in more reliable identification of structural shocks. One leading example is the study by Bernanke, Boivin, and Eliasz (2005; BBE hereafter). One motivation for their paper is the “price puzzle” which is a robust feature of most SVAR studies. The “price puzzle” is a situation where a surprise monetary contraction raises price level in the short run. This is at odds with economic theory. BBE estimate a SFAVAR on 120 macro time series for the US, identify the MP shock using a block-recursive scheme, and compute impulse responses for included variables. They find no evidence of “price puzzle” and conclude that this
empirical pattern is a consequence of misidentified MP shock in VAR.

I use a version of BBE procedure to identify the structural MP shock in the US. Then I use the identified shock as a predictor variable in forecasting exchange rates of OECD currencies out of sample at horizons up to 24 months. Despite delayed overshooting present in most in-sample impulse responses, I do not find any evidence that forecasts with the identified US MP shock outperform RW at any horizons.

Besides the MP shock, the identification procedure I use partially identifies the rest of structural shocks. Specifically, it isolates a group of “fast” shocks that contemporaneously, i.e. within a month, affect mostly financial market variables. These shocks account for a considerable fraction of the in-sample forecast error variance decomposition for exchange rates. Therefore, they may have some forecasting value. I include the “fast” shocks as predictors and show that they do forecast exchange rates of commodity currencies, i.e. the currencies of major commodity exporters in my sample (Australia, Canada, and New Zealand). These “fast” shocks can be interpreted as news about future prospects of the US economy.

The news interpretation of fast shocks is inspired by Beaudry and Portier (2006). Using a two-variable SVAR for total factor productivity (TFP) and stock market index, the authors demonstrate that the shock that contemporaneously affects only stock prices without changing TFP is the same as the shock that shifts TFP in the long run. Their argument is the following. Stock prices are essentially forward-looking. News about anticipated productivity improvements makes investors revise their expectations regarding future dividends. As a result, stock prices, which are discounted sums of expected future dividends, will respond to the news immediately while the change in TFP will happen later.

In the context of my study, news about higher productivity in future makes market participants anticipate a rise in economic activity and, as a result, higher demand for commodities and higher commodity prices when the technological improvement occurs. Sufficiently high storage costs should discourage firms to accumulate inventories and therefore prices of commodities will rise gradually. Exchange rates of commodity exporters are known to be strongly affected by changes in their terms of trade (Chen and Rogoff, 2003). More favorable terms of trade in the future due to higher prices for commodity exports will translate in currency appreciation.

The finding of forecasting value of fast shocks for commodity currencies in this paper is related to a recent study by Chen, Rogoff, and Rossi (2008). These authors document that exchange rates of commodity exporters forecast out of sample prices of commodities exported by those countries better that a RW. From a theoretical perspective, exchange rate can be viewed as a discounted sum of expected values of fundamentals. Change in expectations will result in an immediate change in the exchange rate. Hence, exchange rate should be able to forecast future fundamentals that are unobserved by econometrician but rationally
expected by market participants.

News about future productivity improvement in the US induces a major exchange rate response on impact as well as some non-trivial transition dynamics in subsequent periods caused by a gradual growth in commodity prices. The findings by Chen, Rogoff, and Rossi (2008) are related to the reaction of exchange rate to the news shock on impact. My results, in contrast, are connected to the transition dynamics in the aftermath of the news arrival, specifically, a gradual growth in commodity prices. What allows me successfully forecast prices of commodity currencies at relatively short horizons is the ability of FAVAR to estimate the space of shocks that hit the economy and isolate an empirically relevant group of shocks, i.e. fast shocks, by imposing fairly conventional identifying restrictions.

The rest of the paper proceeds as follows. Section 2 presents FAVAR methodology and identification of structural MP shock. Sections 3 gives details of the forecasting exercise I undertake including forecast performance evaluation. Section 4 describes data. Section 5 reports and discusses empirical results. Section 6 concludes.

2 Factor-augmented vector autoregression

Most forecasting specifications used in the literature are based on cointegration between exchange rate (ER) $s_t$ and its long-run equilibrium level $f_t$ determined by fundamentals:

$$s_{t+h} - s_t = \alpha + \beta (f_t - s_t) + e_{t+h}$$

Different structural macroeconomic models yield different functional forms of $f_t$. For example, the Purchasing Power Parity (PPP) implies that, up to a constant term, $f_t = p_t - p_t^*$ where $p_t$ and $p_t^*$ are logarithms of domestic and foreign price levels respectively. Cointegration implies that exchange rate will converge to its fundamental level in the long-run. This potentially makes cointegrating equations $s_t - f_t$ able to forecast exchange rates at long horizons better than RW. Empirically though the convergence is not very fast, which undermines the forecasting content of cointegrating equations at shorter horizons. According to the literature, forecasting specifications such as 1 are much more successful in forecasting at longer horizons (more than 2-4 years) and perform poorly at shorter horizons (Meese and Rogoff, 1983; Mark, 1995; Mark and Sul, 2001). Occasional positive findings are typically non-robust with respect to the choice of currency or forecast window.

This study does not employ cointegration as a source of forecastability. It takes a different route and, instead, uses structural shocks as predictors. This approach is inspired by “delayed overshooting” effect
widely documented in the literature. A surprise monetary policy tightening in the US leads to nominal appreciation of the US currency in the long run but, in the short run, the US dollar overshoots its new long-run equilibrium level. The peak impulse response is located at horizons 3-16 months (Eichenbaum and Evans, 1995; Faust and Rogers, 2003). MP shocks are shown to be an important driving force of exchange rate movements accounting for 30-40% of forecast error variance decomposition. It follows that if one is able to detect an arrival of MP shock today, she should be able to forecast exchange rate in 3-12 months better than RW. The source of forecasting value of MP shock is “delayed overshooting”. The MP shock would be useless in forecasting if exchange rate impulse responses were flat. The MP-shock-based specification as opposite to cointegration-based specifications will have the form

$$s_{t+h} - s_t = \alpha + \beta \zeta_{t}^{R} + e_{t+h}$$

(2)

where $\zeta_{t}^{R}$ is the MP shock.

One question that naturally arises is how to identify and estimate the MP shock. In principle, it can be done using a structural VAR but the models of this family are notorious for their poor out-of-sample forecasting performance. Furthermore, SVARs are prone to omitted variable bias because of their inability to operate with a sufficiently large number of variables. A leading example is the “price puzzle”, a counter-intuitive empirical finding that a contractionary MP shock raises price level in the short run. A common interpretation of this empirical pattern is that some important variables from the Fed’s information set are omitted and this results in misidentified MP shock. It is conceivable that the Fed tightens MP in an attempt to reduce a spike in future inflation that they anticipate based on available information. The curse of dimensionality prevents an econometrician from including into VAR all variables from the Fed’s information set. Although adding a commodity price index as a forward-looking proxy of inflation expectations removes the price puzzle, there is still no guarantee that the MP shock is identified correctly since other sources of OVB may be present.

Factor Augmented VAR (FAVAR) provides a more profound and systematic treatment of OVB and therefore yields potentially more reliable identification of structural shocks compared with VARs. They are able to accommodate a large number of variables, potentially, all variables from the Fed’s information set. The key assumption that makes estimation feasible is that all macro time series are driven by a small number of common factors/shocks. All interdependencies among the variables work only through the dependence on common factors. The factors are estimated as principal components of data. Effectively, FAVAR turns curse of dimensionality into a blessing: more data help more precisely estimate the space of common shocks.

Bernanke, Boivin, and Eliasz (2005) estimate FAVAR using a panel of 120 US macro time series and identify structural MP shock. They find no evidence of the price puzzle and conclude that it is an artifact of
misspecified VAR.

Factor-based forecasts have a good forecasting record. Stock and Watson (2006) demonstrate their superiority to other forecasting methods with many predictors in forecasting a set of US macro time series. FAVARs proved successful in forecasting inflation at short horizons (Amstad and Potter, 2007).

The rest of this section describes the FAVAR methodology and identification of MP shock in FAVAR. The discussion follows very closely Stock and Watson (2005).

2.1 Factor Augmented Vector Autoregression

Conventional reduced-form VAR can be written as

$$X_t = C(L)X_{t-1} + e_t$$

where $X_t$ is $n \times 1$ vector of endogenous variables and $C(L)$ is a matrix of lag polynomials. If the VAR order is $p$ then the number of parameters to be estimated is $n((2p+1)n-1)/2$ in $C(L)$ and $n(n-1)/2$ in the covariance matrix of $e_t$. VAR is subject to the “curse of dimensionality” since the number of estimated parameters grows as $n^2$. Adding more variables makes estimates less precise. This is why, in practice, VARs rarely include more than 7-12 variables, which represent a very small subset of information available to economic agents and policy makers.

To make the problem manageable, assume that all included variables depend on a relatively small number of common factors:

$$X_{it} = \tilde{\lambda}_i(L)'f_t + u_{it}$$

where $f_t$ is $q \times 1$ vector of dynamic factors, $\tilde{\lambda}_i$ is $q \times 1$ vector of lag polynomials (dynamic factor loadings) of order $k$, and $u_{it}$ is idiosyncratic term (measurement error). Factors $f_t$ are dynamic in the sense that $X_{it}$ depends on contemporaneous and lagged values of $f_t$. Representation 3 assumes that all cross-dependencies among $X_{it}$’s work only through common factors $f_t$. In general, $u_{it}$ is serially correlated:

$$u_{it} = \delta_i(L)u_{it-1} + v_{it}$$

Combining equations (3) and (4) gives

$$X_{it} = \tilde{\lambda}_i(L)'(I - \delta_i(L)L)f_t + \delta_i(L)X_{it-1} + v_{it}$$

$$= \lambda_i(L)'f_t + \delta_i(L)X_{it-1} + v_{it}$$

The evolution of factors is governed by a VAR process:

$$f_t = \Gamma(L)f_{t-1} + \eta_t$$
where \( \eta_t \) is \( q \times 1 \) vector of exogenous shocks that hit the economy with

\[
E(\eta_t \eta_t') = I_q
\]

Shocks and the error term in (5) are assumed uncorrelated

\[
E(\eta_t v_{it}) = 0
\]

as well as the error terms for any two different \( X \)'s

\[
E(v_{it} v_{js}) = 0
\]

for all \( t, s, i, j, i \neq j \).

Introduce the vector of static factors \( F_t \) in such a way that \( X_{it} \) would depend only on contemporaneous values of \( F_t \). Then equation (5) will take the form

\[
X_{it} = \Lambda_i' F_t + \delta_i(L)X_{it-1} + v_{it} \quad (7)
\]

or, in vector notation,

\[
X_t = \Lambda F_t + D(L)X_{t-1} + v_t \quad (8)
\]

where \( \Lambda \) is a matrix of static factor loadings

\[
\Lambda = \begin{pmatrix}
\Lambda_1' \\
\Lambda_2' \\
\vdots \\
\Lambda_n'
\end{pmatrix}
\]

(9)

and

\[
D(L) = \begin{pmatrix}
\delta_1(L) & 0 & \cdots & 0 \\
0 & \delta_2(L) & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \delta_n(L)
\end{pmatrix}
\]

(10)

Comparing dynamic and static factor representations (5) and (8), it is clear that \( F_t \) will be an \( r \times 1 \) vector that contains all elements of \( f_t \) and certain elements of lagged \( f_t \) with \( q \leq r \leq kq \). Equation (6) implies the following law of motion for static factors \( F_t \):

\[
F_t = \Phi(L)F_{t-1} + G\eta_t \quad (11)
\]

Substitute (11) in (8) to obtain

\[
X_t = \Lambda \Phi(L)F_{t-1} + D(L)X_{t-1} + \Lambda G \eta_t + v_t \quad (12)
\]
Equations (11) and (12) constitute what is called *Factor Augmented Vector Autoregression (FAVAR)* model. Equation (11) describes evolution of common factors $F_t$ and equation (12) is a VAR for $X_t$ augmented by lags of $F_t$. If factors $F_t$ are known then estimation of (11)-(12). First, estimate the VAR for factors

$$F_t = \Phi(L)F_{t-1} + \varepsilon_{F_t}$$ (13)

Second, for each $i$, estimate an OLS regression of $X_{it}$ on own lags and lags of factors:

$$X_{it} = \pi_i(L)^tF_{t-1} + \delta_i(L)X_{it-1} + \varepsilon_{X_{it}}$$ (14)

In applications, the number of static factors is typically not greater than 10 and estimation of VAR for factors (13) therefore does not create any difficulty. Given $r$, the number of parameters to be estimated in (14) does not grow with $n$. It follows that, unlike VAR, FAVAR (13)-(14) is not prone to the curse of dimensionality. Instead, the curse is turned to a blessing: adding more variables to FAVAR helps estimate factors more precisely.

### 2.2 Interpretation of factors

#### 2.2.1 Factors as states variables

One way to interpret static factors is to view then as analogs of state variables in the solution of a log-linearized Dynamic Stochastic General Equilibrium Model (DSGE) as suggested by Boivin and Giannoni (2006). The DSGE solution in the state-space form can be written as

$$Y_t = MZ_t$$ (15)

$$Z_t = PZ_{t-1} + Q\eta_t$$ (16)

where $Y_t$ is the vector of endogenous variables and $Z_t$ is the state vector. State variables, components of $Z_t$, include exogenous processes for shocks (e.g., monetary, fiscal, technological) and predetermined endogenous variables (e.g., capital). The components of $Y_t$ refer usually to broad economic concepts (in BBE words) such as production output and tightness of the credit market, which may be unobserved directly or measured with error. In FAVAR, observed time series $X_{it}$ serve as proxies for unobserved concepts $Y_t$. The first FAVAR equation (11) is thus the analog of the state equation (15) while the second one (12) corresponds to the observer equation (16). The measurement errors $u_{it}$ that are present in proxies $X_{it}$ are likely to be serially correlated. Adding lags of $X_{it}$ to the right hand side of the observer equation in FAVAR aims to get rid of that serial correlation.
2.2.2 DFM as parsimonious ARMA-like representation

Another way to interpret factors is to consider the dynamic factor representation of $X_{it}$

$$X_{it} = \lambda_i(L)' f_t + \delta_i(L) X_{it-1} + v_{it} \quad (17)$$

$$f_t = \Gamma(L) f_{t-1} + \eta_t \quad (18)$$

as a multi-shock analog of ARMA model. Indeed, if $f_t = \nu_t$ is a scalar i.i.d. process and $v_{it} \equiv 0$ then equation (17) becomes

$$X_{it} = \lambda_i(L) \nu_t + \delta_i(L) X_{it-1} \quad (19)$$

which characterizes a univariate ARMA process for $X_{it}$.

In general, assume that there are a small number of major economic shocks $\eta_t$ that hit the economy. For simplicity, set $v_{it} \equiv 0$ in (17). In general, each variable $X_{it}$ can be represented as an infinite distributed lag of shocks:

$$X_{it} = a_i(L)' \eta_t \quad (20)$$

One can think about (20) as the solution to a log-linearized DSGE model. Using (18), express $\eta_t$ as

$$\eta_t = (I - \Gamma(L)L) f_t \quad (21)$$

and substitute it in (20) to obtain

$$X_{it} = a_i(L)'(I - \Gamma(L)L) f_t \quad (22)$$

In general, lag polynomial $a_i(L)'(I - \Gamma(L)L)$ is of infinite order but it can be well approximated by a ratio of two finite-order polynomials:

$$X_{it} = \frac{\lambda_i(L)'}{c_i(L)} f_t \quad (23)$$

where, without loss of generality, $c_i(0) = 1$ so that $c_i(L) = 1 - \delta_i(L)L$. A similar trick is used when the infinite MA representation of a univariate time series is approximated by a finite-order ARMA. Multiplying both sides of (23) by $c_i(L)$ and moving lags of $X_{it}$ to the right-hand side yields

$$X_{it} = \lambda_i(L)' f_t + \delta_i(L) LX_{it} \quad (24)$$

The latter expression coincides with equation (17) up to the idiosyncratic error term $v_{it}$. 
2.2.3 Summary

The preceding discussion suggests that factors can be interpreted in several ways. One interpretation relates static factors to a vector of state variable in a DSGE model that describes the economy. Vector $F_t$ consists of exogenous processes for shocks and predetermined endogenous variables. Another potential interpretation suggests considering DFM/FAVAR representation of $X_t$ as a multi-shock analog of ARMA model and dynamic factors as exogenous processes governing economic shocks. Given that I am interested in estimating the space spanned by shocks $\eta_t$ it is not very important perhaps which particular specification I prefers. Unlike factors, shocks $\eta_t$ have clear economic meaning. Strictly speaking, the space of $\eta_t$ spans the space of structural shocks $\zeta_t$. By imposing certain identifying restrictions, I can in principle identify some of the structural shocks. It will be discussed in more detail below.

2.3 Estimation of factors and space of shocks

2.3.1 Static factors

Suppose that the number of factors $r$ is known. Then static factors can be estimated as first $r$ principal components of data. One can extract principal components either from original or prefiltered data. In the former case, factors minimize

$$\frac{1}{nT} \sum_{t=1}^{T} [X_t - \Lambda F_t]' [X_t - \Lambda F_t]$$

with respect to $F_1, ..., F_T$, and $\Lambda$. Technically, factors $F_t$ turn out to be the eigenvectors of sample covariance matrix

$$\frac{1}{nT} \sum_{t=1}^{T} X_t X'_t$$

that correspond its $r$ largest eigenvalues. This method is typically used when when the main objective of a study is forecasting with factors as predictors. If instead one is interested in a more precise estimation of the space of shocks then applying principal components to prefiltered data might be preferable. In this case, factors minimize

$$\frac{1}{nT} \sum_{t=1}^{T} [(I_n - D(L)L)X_t - \Lambda F_t]' [(I_n - D(L)L)X_t - \Lambda F_t]$$

with respect to $F_1, ..., F_T$, $\Lambda$, $D(L)$. Stock and Watson (2005) suggest the following two-stage iteration procedure that implements this estimation.

1. Pick initial $\hat{D}^{(0)}(L)$, compute $\hat{X}_t = (I_n - \hat{D}^{(0)}(L)L)X_t$.
2. Estimate $\hat{F}^{(1)}_t$ as first $r$ principal components of $\hat{X}_t$.
3. Run OLS regression of $X_t$ on $\hat{F}^{(1)}_t$ and own lags, set $\hat{D}^{(1)}(L)$ to estimated coefficients on own lags.
4. Compute $\tilde{X}_t = (I_n - \hat{D}^{(1)}(L)L)X_t$, then estimate $\hat{F}_t^{(2)}$, etc. Iterations will converge soon.

2.3.2 Space of dynamic shocks

Factor-based forecasts typically use static factors as predictors. If the purpose is to compute impulse responses and forecast error variance decompositions to an identified structural shock of interest, then one needs to estimate the space of dynamic shocks $\eta_t$, which, in the DFM context, coincides with the space of dynamic factor innovations. The space of $\eta_t$ spans the space of structural shocks $\zeta_t$.

Suppose that the number of dynamic factors $q$ is known. Re-write the factor equation of FAVAR in the moving average form:

$$F_t = (I - \Phi(L)L)^{-1}\varepsilon_F = (I - \Phi(L)L)^{-1}G\eta_t$$

and substitute it in the FAVAR equation for $X_t$ to obtain

$$X_t = (I - D(L)L)^{-1}\Lambda(I - \Phi(L)L)^{-1}G\eta_t + u_t = A(L)G\eta_t$$

where

$$A(L) = (I - D(L)L)^{-1}\Lambda(I - \Phi(L)L)^{-1}$$

I use a standard normalization:

$$\mathbb{E}(\eta_t\eta_t') = I_q$$

As suggested in Stock and Watson (2005), I choose $r \times q$ matrix $G$ such that maximizes the trace R squared of $A(L)G\eta_t$:

$$\text{tr}(\mathbb{E}(A(L)G\eta_t\eta_t'A(L)'G')) = \text{tr}(G'(\sum_{j=0}^{\infty} A_j'A_j)G)$$

It follows that $G$ that maximizes the trace R squared equals the matrix of eigenvectors of $\sum_{j=0}^{\infty} A_j'A_j$ that correspond to its $q$ largest eigenvalues. Matrix $\sum_{j=0}^{\infty} A_j'A_j$ can be estimated using impulse responses of $X$'s to FAVAR innovations of factors $\varepsilon_F$.

Given $G$ and impulse responses to static factors innovations $A(L)$, I can compute impulse responses to dynamic factor innovations $\eta_t$ as

$$B(L) = A(L)G$$

2.3.3 Number of factors

How do we know the number of static and dynamic factors? In both cases, the answer can be obtained using Bai-Ng information criteria $IC_{p1}$ and $IC_{p2}$ (Bai and Ng, 2002). These criteria are very similar to
well-known Akaike and Schwartz information criteria. Each of them trades off goodness of fit against the number of factors:

$$IC_{p1}(r) = \ln(V(r, \hat{F}^r)) + r\left(\frac{n + T}{n}\right)\ln\left(\frac{nT}{n + T}\right)$$

$$IC_{p2}(r) = \ln(V(r, \hat{F}^r)) + r\left(\frac{n + T}{nT}\right)\ln(C_{nT}^2)$$

where $C_{nT}^2 = \min\{n, T\}$, $r = 1, \ldots, r_{max}$ and

$$V(r, \hat{F}^r) = \frac{1}{nT} \sum_{t=1}^{T} \left[\left(I - \hat{D}^r(L)L\right) X_t - \hat{\Lambda}^r \hat{F}^r_t\right]'\left[\left(I - \hat{D}^r(L)L\right) X_t - \hat{\Lambda}^r \hat{F}^r_t\right].$$

The way to proceed is to choose $r$ that minimizes $IC_{p1}$ and/or $IC_{p2}$. Bai and Ng prove that specific forms of the penalty term, which are different from Akaike and Schwartz, deliver consistent estimates of $r$.

With $r$ known, one can estimate factors $F_t$ as first $r$ principal components of original or prefiltered data and estimate FAVAR as discussed above. Estimated FAVAR yields estimated vector of innovations of data $\varepsilon_{Xt}$ where each element equals the OLS residual of respective $X_{it}$ on own lags and lags of factors. The observer equation (12) of FAVAR suggests that

$$\varepsilon_{Xt} = \Lambda G \eta_t + v_t$$

It is clear that $\varepsilon_{Xt}$ has a factor structure given that $\eta_t$ and $v_t$ are uncorrelated. Each $\varepsilon_{Xit}$ is a linear function of $q$ common factors $\eta_t$ plus idiosyncratic noise $v_{it}$. This suggests that the dimensionality of $\eta_t$ can be determined by applying Bai and Ng information criteria to FAVAR innovations of data $\varepsilon_{Xt}$ (Amengual and Watson, 2007).

### 2.4 Identification of monetary policy shock

Once FAVAR is estimated and mapping $G$ from the space of FAVAR innovations of static factors to the space $\eta_t$ of dynamic shocks is found, one can compute impulse responses and forecast error variance decompositions with respect to $\eta_t$. In general, however, $\eta_t$’s are not true structural shocks but their mixtures:

$$\zeta_t = H \eta_t$$

Rotation matrix $H$ relates reduced-form dynamic factor innovations $\eta_t$ to structural shocks $\zeta_t$. This matrix summarizes our identifying assumptions that are based on a certain economic model. The task is to impose a minimal set of restrictions on elements of $H$ sufficient to identify a row in $H$ that corresponds to a structural shock of interest.

I identify the structural monetary policy shock in the US using Stock and Watson’s (2005) version of the so-called slow/fast scheme suggested by Bernanke, Boivin, and Eliasz (2005). This is a block-recursive
identification scheme. All variables are divided onto three groups: slow-moving series (production indices, labor market variables, consumer prices), the federal funds (FF) rate as an indicator of monetary policy, and fast-moving series (producer prices, expectations, financial market indicators). All structural shocks are also split onto three categories: slow shocks $\zeta^S_t$, the MP shock $\zeta^R_t$, and fast shocks $\zeta^F_t$.

The following identifying assumptions are made:

1. Slow-moving series can respond contemporaneously (i.e. within a month) only to slow shocks. They can respond to the MP shocks and fast shocks only with a lag.

2. The FF rate can respond contemporaneously (within a month) only to the slow shocks and the MP shock. It can respond to fast shocks only with a lag.

3. Fast-moving series can respond contemporaneously to all shocks.

One can notice that the identifying assumptions are similar to those in a well-known block-recursive scheme suggested by Christiano, Eichenbaum, and Evans (1999) for conventional VAR’s, which is considered as state-of-the-art in this class of identification procedures. Motivation of identifying assumptions 1-3 is therefore similar to those in CCE. For example, a surprise MP tightening or bad news about future prospects of the US economy that immediately affects financial market variables is unlikely to switch production cuts and layoffs within one month. This might be harder to motivate though if the study is done on quarterly data.

BBE in fact impose somewhat different identifying restrictions. First of all, they make the FF rate be the observed MP factor. Also, they allow the MP contemporaneously (i.e. within a month) respond to all shocks. This means, in particular, that the Fed can react to developments on the stock market very fast. One can view the two versions of the slow/fast scheme as two extremes. The BBE version assumes that Fed can process new information very quickly and react immediately. The version by Stock and Watson (2005), which I use in this study, reflects a more conservative view about information capacity and responsiveness of Fed.

Let $B_j^\ast$ be $n \times q$ matrices of impulse responses to structural shocks $\zeta_t$:

$$X_t = B^\ast(L)\zeta_t = \sum_{j=0}^{\infty} B_j^\ast \zeta_{t-j}$$

Then assumptions 1-3 of the slow/fast scheme can be summarized as:

$$\varepsilon_{Xt}^{slow} = B_0^{slow,S} \zeta^S_t + v_t^{slow}$$

$$\varepsilon_{Xt}^{FF} = B_0^{FF,S} \zeta^S_t + B_0^{FF,R} \zeta^R_t + v_t^{FF}$$

$$\varepsilon_{Xt}^{fast} = B_0^{fast,S} \zeta^S_t + B_0^{fast,R} \zeta^R_t + B_0^{fast,F} \zeta^F_t + v_t^{fast}$$
In practice, one needs to know the number of slow dynamic shocks $q^S$. It can be obtained by using Amengual and Watson (2007) method applied to slow-moving series only. Equation (26) implies that FAVAR innovations of slow-moving series depend on common factors $\zeta_t^S$. Hence, applying Bai-Ng criterion to $\varepsilon_{Xt}^{slow}$ will yield $q^S$.

Given $q^S$, one runs a reduced-rank regression of $\varepsilon_{Xt}^{slow}$ on the space of factor innovations $\eta_t$ imposing a restriction that the column rank of the coefficient matrix equals $q^S$. Indeed, $\eta_t$ spans the space of $\zeta_t$ and they are related through a rotation matrix $H$. Hence, $\zeta_t^S$ that are common factors for $\varepsilon_{Xt}^{slow}$ have to be $q^S$ orthogonal linear combinations of $\eta_t$. This is exactly what RR regression of $\varepsilon_{Xt}^{slow}$ on $\eta_t$ delivers:

$$\varepsilon_{Xt}^{slow} = B_0^{slow,S}H^S\eta_t + v_t^{slow}$$

where $H^S$ is $q^S \times q$ submatrix of $H$:

$$H = \begin{pmatrix} H^S \\ H^R \\ H^F \end{pmatrix}$$

Once the subspace of slow shocks $\zeta_t^S$ is estimated, one can exploit equation (27) to identify the MP shock. It equals the residual from OLS regression where the dependent variable is OLS projection of FAVAR innovation of Federal Funds Rate $\varepsilon_{Xt}^{FF}$ on the space of factor innovations $\eta_t$ and regressors are slow shocks $\zeta_t^S$. The knowledge of the space of $\zeta_t^S$ (as opposite to its components) is just enough to identify the MP shock $\zeta_t^R$.

Finally, the space of fast shocks $\zeta_t^F$ is estimated as the subspace of $\eta_t$ orthogonal to $\zeta_t^S$ and $\zeta_t^R$.

To summarize, the slow/fast scheme yields three outputs: exactly identified MP shock, the space of slow shocks and the space of fast shocks. The MP shock and delayed overshooting produced by it as reported by a number of SVAR studies is the main motivation of the main idea of this paper to use structural shocks as predictors. Nevertheless, decomposition of the rest of shocks into two groups, slow and fast, may also be helpful despite the fact that the procedure does not isolate individual structural shocks within each of the two groups. First, it can improve forecast performance by discarding the group of shocks that accounts for a negligible fraction of forecast error variance decomposition. Second, finding that a given group of shocks forecasts exchange rates out-of-sample may provide useful feedback to theory and policy advice.

### 3 Exchange rate forecasting

My general forecasting equation has the form

$$s_{t+h} - s_t = \beta'\zeta_t^g + \epsilon_{t+h}$$
where $s_t$ is log bilateral nominal exchange rate with the dollar, $\zeta^g_t$ is a group of partially identified structural shocks $g$, a subvector of $\zeta_t$. Specifically, I consider five forecast specifications each corresponding to a particular group $\zeta^g_t$:

1. all shocks: $\zeta^g_t = (\zeta^{S}_t, \zeta^{R}_t, \zeta^{F}_t)'$;
2. slow shocks: $\zeta^g_t = \zeta^{S}_t$;
3. MP shock: $\zeta^g_t = \zeta^{R}_t$;
4. fast shocks: $\zeta^g_t = \zeta^{F}_t$;
5. MP and fast shocks: $\zeta^g_t = (\zeta^{R}_t, \zeta^{F}_t)'$.

The purpose of trying different combinations of shocks as predictors is two-fold. First, omitting empirically irrelevant regressions should reduces the degree of sample uncertainty and thus raise the quality of forecast. Second, it potentially provides a means of testing competing theories of exchange rate behavior. If, for example one theory stresses the importance of monetary policy shocks and predicts a non-trivial impulse response, e.g. with a regular (Dornbusch, 1976) or delayed overshooting, then such a theory can be tested by an out-of-sample forecast with the identified MP shock as predictor. In general, ability to forecast out-of-sample better than a random walk has been considered by the literature started by Meese and Rogoff (1983) as the most appealing and, at the same time, the hardest to obtain empirical evidence in favor of a structural exchange rate model. Of course, valid identification of structural shocks of interest is critical: a theory under consideration can be right but misidentified structural shock will result in the theory being rejected. For the latter reason, I include specification 5 in to the list above. If the slow/fast procedure misidentifies the US MP shock then the true shock is likely to be a linear combination of misidentified MP and fast shocks.

Following the literature, I evaluate forecast performance by using Theil’s U statistic (TU). It is defined as the ratio of root mean squared forecast errors of the model’s and RW forecasts:

\[
TU = \frac{RMSFE}{RMSFE^{RW}}
\]

where

\[
RMSFE = \sqrt{\sum_{t=P+1}^{T} (s_t - \hat{s}_t^h)^2}
\]

\[
RMSFE^{RW} = \sqrt{\sum_{t=P+1}^{T} (s_t - s_{t-h})^2}
\]

\(s_t\) is actual log nominal exchange rate, \(\hat{s}_t^h\) is the forecast produced by the model at date \(t - h\), \(s_{t-h}\) random walk forecast from the same date’s perspective, and assuming that the first forecast is done for date \(P + 1\). Here, sample interval \(t \in [P + 1, T]\) is called forecast window.

It is well known that forecast performance of many structural exchange rate models in the literature is not robust with respect to the choice of forecast window, a sample time interval on which TU or other evaluation statistic computed. A model may beat a RW one the sample of the seventies and eighties but perform poorly once nineties are added to the sample. To check for the robustness of a forecast with respect to the forecast window, I follow Rogoff and Stavrakeva (2008) and look at the recursive TU statistics, i.e. TU statistics computed for different dates of the first forecast. A forecast outperforms a RW robustly with respect to the choice of forecast window if its TU remains significantly less than one for all (or majority) of dates of first forecast.

The statistical significance of TU statistics is evaluated by bootstrap method similar to what Rogoff and Stavrakeva (2008) use in their study. Under the null hypothesis, the nominal exchange rate follows a driftless random walk. This implies that its monthly increments are i.i.d. The estimated dynamic shocks \(\zeta_t\) that serve as predictors are also i.i.d. If the number of variables \(n\) included into FAVAR is large enough then, according to Stock and Watson (2002), one can treat estimated static factors as observables. Hence, \(\zeta\)’s that are rotated residuals from a VAR for factors can be handled as “regular” residuals of an observed variable. At each round \(l\) of resampling, I draw from \((\Delta s_t, \zeta_t)\) to get \((\Delta s_t^{*l}, \zeta_t^{*l})\), \(t = 1, 2, \ldots, T\). Then I construct the simulated series of NER as

\[
s_t^{*l} = \sum_{k=0}^{t-1} \Delta s_t^{*l} - k
\]

from where I obtained simulated series of \(h\)-differences \(\Delta_h s_t^{*l} = s_t^{*l} - s_{t-h}^{*l}\). I forecast \(\Delta_h s_{t+h}^{*l}\) with \(\zeta_t^{*l}\) as predictors and save forecast values \(\Delta_h \hat{s}_{t+h}^{*l}\). For each date of first forecast \(P\), I compute the Theil’s statistic \(TU_h^{*l}(P)\). I repeat the entire process \(L\) times. As a result, for each horizon \(h\) and each date of
the first forecast $P$, I have a bootstrapped distribution $\{TU^*_h(P)\}_{l=1}^L$. Under the RW null, $TU_h(P) = 0$. The bootstrapped p-value of $TU_h(P)$ is computed as a fraction of values $TU^*_h(P)$ of the bootstrapped distribution below $TU_h(P)$:

$$p_h(P) = \mathbb{P}[TU^*_h(P) < TU_h(P)]$$

I follow the literature and use the significance level of 10%. I consider a forecast to be robust with respect to the choice of forecast window if $p_h(P) \leq 0.1$ over the most of the range of $P$. The earliest first date of forecast in this study is 1983M1 and the latest 2000M12 to make sure that the forecast window is no shorter than 60 months.

Figure 1 gives an example of good and robust forecast. For all possible dates of first forecast 1983M1-2000M12, the bootstrapped p-value is below the critical level of 0.10 indicating that the forecast of CAD/USD 20 months ahead with all shocks as predictors robustly outperforms a random walk.

Figure 2 gives an example of a poor forecast. For all forecast windows, the bootstrapped p-value is located above 0.10 suggesting that the forecast of CAD/USD 1 month ahead with all shocks as predictors is consistently inferior to a RW.

Figure 3 shows a forecast that outperforms a RW if one evaluates its performance based on the full sample 1983M1-2006M12. But this forecast does worse than a RW if one evaluates it, say, on 1994M1-2006M12 forecast window. This forecast is thus not robust with respect to the choice of forecast window.

The forecast shown on figure 4 is not uniformly better than a RW for all forecast windows. Specifically, there are two spikes in the bootstrapped p-value, one in 1988 and the other in 1997. Except these two episodes of temporary deterioration in performance, the forecast with shocks beats a RW. I classify a forecast as fairly robust if its bootstrapped p-value remains under the 0.1 level for the most of the range of forecast windows. Figure 4 is an example of such a forecast.

4 Data

I use data from several sources, Stock and Watson (2005, 2007), FRED, OECD, and Global Financial Data. All data are at monthly frequencies. The dataset includes the US and other 9 OECD countries. It covers the time period 1973M1-2006M12.

I run a separate forecasting exercise for each country paired with the US. For example, to forecast the CAD/USD, I take data for Canada and the US and augment them with world commodity prices. I estimate FAVAR on such a bilateral dataset, identify shocks and use them as predictors in forecasting CAD/USD. For each country, I have a number of macro time series and financial market variables.

The US macro data are borrowed from Stock and Watson (2005, 2007). A few variables were added
Figure 1: Good and robust forecast

Figure 2: Poor forecast
Figure 3: Good but non-robust forecast

Figure 4: Good and fairly robust forecast
from FRED. In total, 108 variables are included and they cover data on income and consumption, production, employment, housing, orders and inventories, confidence indicators, and prices.

The data for non-US OECD countries in my sample come from the Main Economic Indicators (MEI) database produced by OECD. The included countries are Australia, Canada, Denmark, France, Germany, Iceland, Italy, Japan, Norway, New Zealand, Sweden, Switzerland, and UK. Given that this study focuses on exchange rate behavior, I consider all OECD countries that had relatively flexible exchange rate regime based on Reinhart and Rogoff’s (2004) classification. Korea, Mexico, and Turkey do not pass this criterion as they featured multiple switches between fix to float regimes over the sample period. I do not try to forecast Euro/USD because of the short lifespan of the Euro. I do not consider any countries from the Eurozone for that reason.

Financial time series and real commodity prices were taken from the Global Financial Database. The financial variables are yields on government and private fixed-income securities for different maturities as well as stock market indicators and exchange rates. Term and default spreads were computed from relevant interest rate series (yields). The values of all financial variables are end-of-month.

Included commodities are aluminum, cocoa, coffee, copper, corn, cotton, gold, heating oil, lead, live cattle, live hogs, nickel, oil Brent, oil Dubai, oil West Texas Instrumental, silver, soybeans, sugar, wheat, zinc, oil (average price), metals (real price index, IMF), agricultural raw materials (real price index, IMF), beverages (real price index, IMF), food (real price index, IMF). Adding commodity prices to the dataset aims to control for global aggregate demand conditions.

The number of available non-US series per country differs much across countries, from 92 for Japan to about 30 for New Zealand. The composition of the data for non-US countries is roughly similar to that for the US except the number of available series per country being less in general. It means, in particular, that the numbers of slow- and fast-moving series per country are approximately equal.

All non-financial time series are seasonally adjusted. When estimating factors, all original time series were transformed to obtain stationarity, standardized, and adjusted to outliers.

5 Empirical findings

5.1 Structural shocks as potentially good predictors: in-sample SFAVAR evidence

This study is motivated by what is called “delayed overshooting”. Eichenbaum and Evans (1995) identify the MP shock in a six-variable VAR. They show that the impulse response of exchange rates to the MP shock is different from what the Dornbusch (1976) model predicts. Following a surprise MP tightening in the US, the dollar appreciates on impact and keeps appreciating further. The peak effect happens in 1 to 3
years after which the US currency starts depreciating and converges to its new long-run equilibrium level.

A number of studies challenge the delayed overshooting finding. Kilian (1998) shows that its statistical significance as reported in Eichenbaum and Evans (1995) might be due to the failure to compute confidence bands with correct coverage probability. He shows that if one uses bootstrapped confidence bands instead of asymptotic ones as EE do then the delayed overshooting becomes statistically insignificant.

Faust and Rogers (2003) relax contemporaneous identifying restrictions imposed by EE by removing what they call dubious restrictions, i.e. those that do not have a strong support by economic theory. Their procedure yields a set identification. The authors claim that the set-identified impulse responses suggest that delayed overshooting may not be a real-life phenomenon but instead a consequence of the dubious assumptions.

Scholl and Uhlig (2007) revisit EE using sign restrictions and re-confirm delayed overshooting. Furthermore, they demonstrate that imposing monotone impulse response of exchange rate to MP shock as predicted by the Dornbusch model yields unconventional and counterintuitive impulse responses for other variables.

All the above mentioned studies are in-sample. In this study, I approach the delayed overshooting from the out-of-sample perspective. If this is an economically significant real-life phenomenon, the identified MP shock should be able to produce non-trivial forecasts of the exchange rate, at least, at horizons close to the peak effect.

Figures 5 to 8 show impulse responses of exchange rates to identified MP shock in a FAVAR for the US. Here, I use Stock and Watson’s (2005) data and their version of BBE’s slow/fast identification procedure. The estimation period is 1973M1-2003M12. Although delayed overshooting is not as well pronounced as reported in earlier studies such as EE, for example (in the case of Canada there is no overshooting at all), there might be still some potential for forecast improvement over a random walk. One can notice that, in all for cases, the US dollar appreciates during first several months after the shock. This means that if I am able to detect arrival of a contractionary MP shock at date $t$, I can predict the exchange rate to be appreciating at date $t + 1$, $t + 2$, etc. This forecast can be potentially better than one by a RW.

Tables 1 to 4 show forecast error variance decompositions of four exchange rates at various horizons obtained from the same structural FAVAR. They suggest that MP shocks account for a remarkable fraction of exchange rate variability. This observation reassures that the MP shock is likely to have a non-trivial forecasting content with respect to exchange rates.

A similar reasoning can be applied to other shocks, in principle. As it was already discussed, the slow/fast identification scheme partially identifies two other groups of shocks in addition to the MP shock.
Figure 5: Impulse responses of CAD/USD to US MP shock. Dashed and dotted lines show 67 and 90% bootstrapped confidence bands respectively.

Figure 6: Impulse responses of CHF/USD to US MP shock. Dashed and dotted lines show 67 and 90% bootstrapped confidence bands respectively.
Figure 7: **Impulse responses of GBP/USD to US MP shock.** Dashed and dotted lines show 67 and 90% bootstrapped confidence bands respectively.

Figure 8: **Impulse responses of JPY/USD to US MP shock.** Dashed and dotted lines show 67 and 90% bootstrapped confidence bands respectively.
Table 1: **Forecast error variance decomposition for CAD/USD.**

<table>
<thead>
<tr>
<th>Horizon, months</th>
<th>0</th>
<th>2</th>
<th>4</th>
<th>8</th>
<th>16</th>
<th>24</th>
</tr>
</thead>
<tbody>
<tr>
<td>Idiosyncratic</td>
<td>76.8</td>
<td>73.7</td>
<td>73.9</td>
<td>74.4</td>
<td>74.3</td>
<td>74.2</td>
</tr>
<tr>
<td>Slow</td>
<td>3.0</td>
<td>2.5</td>
<td>3.0</td>
<td>2.8</td>
<td>2.5</td>
<td>2.4</td>
</tr>
<tr>
<td>Monetary</td>
<td>8.3</td>
<td>11.7</td>
<td>12.4</td>
<td>12.9</td>
<td>13.3</td>
<td>13.5</td>
</tr>
<tr>
<td>Fast</td>
<td>11.8</td>
<td>12.0</td>
<td>10.6</td>
<td>10.0</td>
<td>9.9</td>
<td>9.9</td>
</tr>
</tbody>
</table>

Table 2: **Forecast error variance decomposition for CHF/USD.**

<table>
<thead>
<tr>
<th>Horizon, months</th>
<th>0</th>
<th>2</th>
<th>4</th>
<th>8</th>
<th>16</th>
<th>24</th>
</tr>
</thead>
<tbody>
<tr>
<td>Idiosyncratic</td>
<td>32.0</td>
<td>32.5</td>
<td>33.0</td>
<td>33.7</td>
<td>34.1</td>
<td>34.2</td>
</tr>
<tr>
<td>Slow</td>
<td>8.2</td>
<td>6.1</td>
<td>6.4</td>
<td>6.0</td>
<td>6.0</td>
<td>6.1</td>
</tr>
<tr>
<td>Monetary</td>
<td>43.9</td>
<td>46.9</td>
<td>44.9</td>
<td>43.8</td>
<td>43.3</td>
<td>42.9</td>
</tr>
<tr>
<td>Fast</td>
<td>15.8</td>
<td>14.4</td>
<td>15.7</td>
<td>16.4</td>
<td>16.7</td>
<td>16.8</td>
</tr>
</tbody>
</table>

Specifically, it estimates the subspaces of slow and fast shocks respectively. As a reminder, slow shocks constitute the only driving force of slow-moving series such as output and employment within a month. Fast shocks affect within a month only fast-moving series such as financial market variables and expectations. Inspecting Tables 1 to 4, one can notice that fast shocks are empirically important for exchange rates since they have a considerable fraction in exchange rate variance decompositions. Slow shocks, on the other hand, explain a rather moderate fraction of exchange rate variability.

A large fraction in variance decomposition by no means guarantees that fast shock will be good predictors of exchange rates. For example, the impulse responses to fast shocks may look like a random walk, i.e. when a fast shock shifts the exchange rate to a new long-run level on impact and there is no any transitional dynamics. One could check whether this is the case or not by inspecting impulse responses to individual fast shocks. The slow/fast scheme, however, does not identify individual fast shocks but only all fast shocks as a group making such a verification unfeasible.

Whether fast shocks or any other group of shock are good predictors or not is an empirical matter anyway. In this study, I test the ability of all three groups of shocks to forecast exchange rates out-of-sample directly.

Table 3: **Forecast error variance decomposition for GBP/USD.**

<table>
<thead>
<tr>
<th>Horizon, months</th>
<th>0</th>
<th>2</th>
<th>4</th>
<th>8</th>
<th>16</th>
<th>24</th>
</tr>
</thead>
<tbody>
<tr>
<td>Idiosyncratic</td>
<td>49.6</td>
<td>52.6</td>
<td>53.5</td>
<td>54.5</td>
<td>55.0</td>
<td>55.2</td>
</tr>
<tr>
<td>Slow</td>
<td>8.0</td>
<td>4.9</td>
<td>4.6</td>
<td>3.9</td>
<td>3.7</td>
<td>3.8</td>
</tr>
<tr>
<td>Monetary</td>
<td>29.1</td>
<td>31.6</td>
<td>30.1</td>
<td>29.4</td>
<td>29.1</td>
<td>28.8</td>
</tr>
<tr>
<td>Fast</td>
<td>13.3</td>
<td>10.9</td>
<td>11.7</td>
<td>12.2</td>
<td>12.3</td>
<td>12.3</td>
</tr>
</tbody>
</table>

Table 3: **Forecast error variance decomposition for GBP/USD.**
Table 4: **Forecast error variance decomposition for JPY/USD.**

<table>
<thead>
<tr>
<th>Horizon, months</th>
<th>0</th>
<th>2</th>
<th>4</th>
<th>8</th>
<th>16</th>
<th>24</th>
</tr>
</thead>
<tbody>
<tr>
<td>Idiosyncratic</td>
<td>52.1</td>
<td>55.7</td>
<td>56.6</td>
<td>56.9</td>
<td>57.1</td>
<td>57.1</td>
</tr>
<tr>
<td>Slow</td>
<td>6.3</td>
<td>2.5</td>
<td>1.8</td>
<td>1.1</td>
<td>0.7</td>
<td>0.6</td>
</tr>
<tr>
<td>Monetary</td>
<td>29.7</td>
<td>31.6</td>
<td>30.5</td>
<td>30.2</td>
<td>30.4</td>
<td>30.3</td>
</tr>
<tr>
<td>Fast</td>
<td>11.8</td>
<td>10.2</td>
<td>11.1</td>
<td>11.8</td>
<td>11.8</td>
<td>11.8</td>
</tr>
</tbody>
</table>

Table 5: **Performance of factor-based forecasts relative to driftless random walk.** Each row corresponds to a specific set of predictors as indicated in the first column. Each column starting from the second one corresponds to a currency whose exchange rate is forecast. Each cell shows horizons at which respective factor-based forecasts robustly (with respect to forecast window) outperforms RW. First estimation period 1973M1; first forecast period 1983M1; last forecast period 2006M12.

### 5.2 Structural shocks as out-of-sample predictors

Tables 5 and 6 show the out-of-sample performance of forecasting models with shocks as predictors. One can make several observations. First, no single group of shocks forecasts exchange rates uniformly well. Furthermore, neither slow shocks as a group nor MP shock appear to be helpful in forecasting exchange rates out-of-sample better than RW. The group of fast shocks have forecasting value with respect to several currencies. Second, forecastability is rather poor in general. Third, there is a group of currencies that seem to be forecast relatively more systematically and better compared with the rest. These are Australia, Canada, and New Zealand. It is worth noting that fast shocks have good forecasting content with respect to these three exchange rate at horizons 1-2 years. What the three countries have in common is that all of them are major commodity exporters. Starting from late 1980s – early 1990s, fast shocks also forecast three other prices of three other currencies from my sample, Denmark, UK, and Sweden.

### 5.3 What makes commodity currencies easier to forecast?

As documented in the previous subsection, commodity currencies seem to make a success story in the sense that they are relatively better forecast compared with other currencies. It is fast shocks that help forecast these three exchange rates at horizons 1 to 2 years. One possible interpretation is that current and future commodity prices constitute the most important fundamental variable for the commodity currencies. It is assumed that the commodities exported by these countries are not storable. Fast shocks can be interpreted as
Table 6: Performance of factor-based forecasts relative to driftless random walk. Each row corresponds to a specific set of predictors as indicated in the first column. Each column starting from the second one corresponds to a currency whose exchange rate is forecast. Each cell shows horizons at which respective factor-based forecasts robustly (with respect to forecast window) outperforms RW. First estimation period 1973M1; first forecast period 1983M1; last forecast period 2006M12.

<table>
<thead>
<tr>
<th>Predictors</th>
<th>JPN</th>
<th>NOR</th>
<th>NZL</th>
<th>SWE</th>
</tr>
</thead>
<tbody>
<tr>
<td>All shocks</td>
<td>none</td>
<td>none</td>
<td>none</td>
<td>none</td>
</tr>
<tr>
<td>Slow shocks</td>
<td>none</td>
<td>none</td>
<td>none</td>
<td>none</td>
</tr>
<tr>
<td>MP shock</td>
<td>none</td>
<td>none</td>
<td>none</td>
<td>none</td>
</tr>
<tr>
<td>Fast shocks</td>
<td>none</td>
<td>none</td>
<td>none</td>
<td>none</td>
</tr>
<tr>
<td>MP + fast shocks</td>
<td>none</td>
<td>none</td>
<td>none</td>
<td>none</td>
</tr>
</tbody>
</table>

Table: $a$ - after 1988; $c$ - after 1992

news about future prospects of the US or world economy. Such news affects financial market variables but does not influence output and employment within one month. Good news about future productivity growth will drive stock prices up. In several months, when the productivity improvement materializes, the increased demand for commodities will raise their prices and, as a result, exchange rates of commodity exporters will appreciate. Thus detecting an arrival of the fast shock will help forecast the future value of a commodity currency through anticipated increase in commodity prices.

Figure 9: Anticipated rise in fundamental. At $t = T$, it becomes known that fundamental will rise from $f$ to $f'$ at $t = T'$

The interpretation of fast shocks as news about future productivity receives support from forecast error variance decompositions for stock market indicators shown in tables 7-10. One can see that fast shocks
### Table 7: Forecast error variance decomposition for S&P500.

<table>
<thead>
<tr>
<th>Horizon, months</th>
<th>0</th>
<th>2</th>
<th>4</th>
<th>8</th>
<th>16</th>
<th>24</th>
</tr>
</thead>
<tbody>
<tr>
<td>Idiosyncratic</td>
<td>10.4</td>
<td>5.9</td>
<td>5.8</td>
<td>5.9</td>
<td>6.0</td>
<td>6.0</td>
</tr>
<tr>
<td>Slow</td>
<td>5.7</td>
<td>10.9</td>
<td>11.5</td>
<td>12.2</td>
<td>12.7</td>
<td>12.9</td>
</tr>
<tr>
<td>Monetary</td>
<td>0.1</td>
<td>1.3</td>
<td>1.7</td>
<td>2.0</td>
<td>2.1</td>
<td>2.1</td>
</tr>
<tr>
<td>Fast</td>
<td>83.8</td>
<td>82.0</td>
<td>81.0</td>
<td>79.9</td>
<td>79.2</td>
<td>79.0</td>
</tr>
</tbody>
</table>

### Table 8: Forecast error variance decomposition for S&P industrials.

<table>
<thead>
<tr>
<th>Horizon, months</th>
<th>0</th>
<th>2</th>
<th>4</th>
<th>8</th>
<th>16</th>
<th>24</th>
</tr>
</thead>
<tbody>
<tr>
<td>Idiosyncratic</td>
<td>11.2</td>
<td>6.3</td>
<td>6.2</td>
<td>6.2</td>
<td>6.3</td>
<td>6.4</td>
</tr>
<tr>
<td>Slow</td>
<td>6.3</td>
<td>11.5</td>
<td>12.0</td>
<td>12.7</td>
<td>13.1</td>
<td>13.3</td>
</tr>
<tr>
<td>Monetary</td>
<td>0.0</td>
<td>0.9</td>
<td>1.3</td>
<td>1.7</td>
<td>1.7</td>
<td>1.7</td>
</tr>
<tr>
<td>Fast</td>
<td>82.4</td>
<td>81.2</td>
<td>80.5</td>
<td>79.4</td>
<td>78.8</td>
<td>78.6</td>
</tr>
</tbody>
</table>

### Table 9: Forecast error variance decomposition for S&P dividend yield.

<table>
<thead>
<tr>
<th>Horizon, months</th>
<th>0</th>
<th>2</th>
<th>4</th>
<th>8</th>
<th>16</th>
<th>24</th>
</tr>
</thead>
<tbody>
<tr>
<td>Idiosyncratic</td>
<td>21.8</td>
<td>12.8</td>
<td>12.5</td>
<td>12.7</td>
<td>13.1</td>
<td>13.3</td>
</tr>
<tr>
<td>Slow</td>
<td>7.4</td>
<td>13.6</td>
<td>15.9</td>
<td>18.6</td>
<td>21.8</td>
<td>23.2</td>
</tr>
<tr>
<td>Monetary</td>
<td>0.5</td>
<td>2.1</td>
<td>2.5</td>
<td>2.7</td>
<td>2.5</td>
<td>2.3</td>
</tr>
<tr>
<td>Fast</td>
<td>70.4</td>
<td>71.5</td>
<td>69.1</td>
<td>65.9</td>
<td>62.6</td>
<td>61.2</td>
</tr>
</tbody>
</table>

### Table 10: Forecast error variance decomposition for S&P P/E ratio.

<table>
<thead>
<tr>
<th>Horizon, months</th>
<th>0</th>
<th>2</th>
<th>4</th>
<th>8</th>
<th>16</th>
<th>24</th>
</tr>
</thead>
<tbody>
<tr>
<td>Idiosyncratic</td>
<td>54.2</td>
<td>39.2</td>
<td>38.9</td>
<td>39.1</td>
<td>39.6</td>
<td>39.8</td>
</tr>
<tr>
<td>Slow</td>
<td>4.7</td>
<td>9.0</td>
<td>10.9</td>
<td>13.6</td>
<td>17.1</td>
<td>18.7</td>
</tr>
<tr>
<td>Monetary</td>
<td>0.0</td>
<td>0.3</td>
<td>0.3</td>
<td>0.3</td>
<td>0.2</td>
<td>0.1</td>
</tr>
<tr>
<td>Fast</td>
<td>41.1</td>
<td>51.5</td>
<td>49.9</td>
<td>47.0</td>
<td>43.2</td>
<td>41.4</td>
</tr>
</tbody>
</table>
account for up to 80% of variation of these variables. Beaudry and Portier (2006) consider a two-variable VAR with the stock market index and TFP and apply two alternative identification procedures, one at a time. First, they identify the productivity shock through the long-run run restriction that only this shock affects TFP in the long run. Second, the identify what they call the news shock by imposing a restriction that it is contemporaneously orthogonal to the VAR innovation in TFP. Beaudry and Portier (2006) compare the impulse responses to the productivity shock identified by the first scheme with impulse responses to the news shock identified by the second scheme. The two sets of impulse responses turn out to be remarkably close to each other. This gives rise to the interpretation that news shock brings information about future changes in productivity.

The news shock affects mostly financial market variables at short horizons. Later, when positive news materializes into an improvement in productivity, increased world demand for commodities will drive up their prices and improve terms of trade of commodity exporters. This will result in appreciation of their currencies. Figure 9 provides a graphical illustration.

6 Concluding remarks

In this study, I forecast exchange rates out of sample using partially identified structural shocks as predictors. The using of a relevant set of partially identified structural shocks improves forecast performance. Fast shocks partially identified by a version of the fast/slow scheme help forecast exchange rates for commodity currencies at horizons 4 to 24 months. One possible interpretation is that news about future productivity growth in the US makes one anticipate higher commodity prices and, as a result, appreciation of commodity currencies in the future. There is also an exchange rate response to the news on impact, which provides forecasting content to commodity exporters’ exchange rates with respect to future commodity prices as documented in Chen, Rogoff, and Rossi (2008).

It proves helpful to preselect predictors by inspecting in-sample impulse responses and variance decompositions. For example, for Australia, using all shocks as predictors results in a poor forecast while keeping only fast shocks improves it over a RW at a number of horizons. Delayed overshooting does not find any empirical support from out-of-sample forecasting perspective: identified MP shock does not have any forecasting value for all currencies. Except commodity currencies, identified structural shocks do not help forecast exchange rates of non-Euro OECD countries at horizons up to 2 years.
References


