Exchange Rate Direction of Change Forecasting and Divisia Monetary Aggregates

Jane Binner, Luis Molinas Sosa, Megan Hart, Peter Tino

Department of Financial Inclusion
Central Bank of Paraguay

September, 2023

Motivation

- Divisia monetary aggregates have been shown to be an improvement on the simple-sum monetary aggregates.
- One problem they may help with is forecasting exchange rates and their direction of change.
- Meese and Rogoff (1983) convincingly argued that no model could outperform a driftless random walk in predicting exchange rates (or their direction of change).
- Many studies have tried to find a model that improves on the random walk but results have been mixed.
- The use of machine learning techniques and Divisia monetary aggregates and User Cost Prices can help in this matter as well.
- In particular, we we will look at the Generalized Matrix Linear Vector Quantization (GMLVQ) algorithm.

- *What we want to know*:
  1. Can the GMLVQ algorithm outperform a benchmark model in forecasting exchange rate direction, irrespective of whether we use simple-sum aggregates and reference interest rates or Divisia indexes and User Cost Prices?
  2. Do Divisia indexes and User Cost Prices outperform simple-sum aggregates and reference interest rates when forecasting the direction of change of exchange rates?

Objectives

- This study includes the User Cost Price and Divisia monetary aggregates as variables in the GMLVQ algorithm to see if forecasting power improves on a benchmark model when the aforementioned variables are compared with a reference interest rate and simple sum monetary aggregates (respectively), along other relevant macroeconomic variables.
- Specifically, this study looks at the USD/EUR, MXN/EUR, CNY/EUR, USD/MXN, USD/CNY exchange rates.

Preview of Results

- *This study finds that:*
- The GMLVQ algorithm outperforms the benchmark model in almost every forecasting horizon.
- The use of the User Cost Price and Divisia aggregates tends to improve accuracy in the GMLVQ algorithm vs the reference interest rate and Simple-sum aggregates.
- The User Cost Price and Divisia are picked as more relevant when directly compared to the reference interest rate and Simple-sum aggregates.

Previous Literature

- Purchasing Power Parity (PPP) and UIP analyses and discussions can be found as far back as the sixties (see, for instance, Balassa (1964)). Dornbusch (1976) proposed a Sticky Price (SP) model based on monetary fundamentals and Frankel (1979) further developed this framework by emphasizing the role of expectations. Hooper and Morton (1982) extended this model to include current account balances.
- However, Meese and Rogoff (1983) wrote a seminal study in which they convincingly argued that no exchange rate model can outperform a driftless random walk in out-of-sample forecasting. Since then, Mark (1995) proposed that at longer horizons a monetary fundamentals model could provide with better out-of-sample forecasts. This model has been subject to criticism by Killian (1999) and Faust et al (2003) where they argue that improvements occur only with a two-year window and disappear afterwards.

- More recent attempts which have shown more promising results. Lothian and Wu (2011) show that UIP has remarkable forecasting power in longer time horizons. Wright (2008) shows that Bayesian Model Averaging outperforms the random walk in shorter time horizons. Lace et al. (2015) argue that the EUR/USD exchange rate can be determined by government yields in the short-run.

Previous Literature

- Bissoondeeal et al. (2008) present evidence that neural network models typically perform better than linear models and traditional nonlinear models in forecasting exchange rates.
- Further, Radityo, Munajat and Budi (2017) were able to employ various neural network methods to forecast cryptocurrency trade.
- Galeshchuk and Mukherjee (2017) successfully employed a convolutional neural network to predict changes in the direction of exchange rates.

- Barnett and Kwag's (2006) were able to show that the use of Divisia monetary aggregates and the User Cost Price dramatically improve the forecasting power of structural models. In a similar vein, Ghosh and Bhadury (2018) show that Divisia Monetary aggregates are powerful indicators of exchange rate movements for several economies. Molinas Sosa, Binner and Tong (2021) extended the Barnett and Kwag methodology for exchange rates that involved negative interest rates, and found that models including Divisia aggregates and the User Cost Price demonstrated improved long-term forecasting ability.

- The User Cost Price and Divisia Monetary aggregates were derived by Barnett (1978, 1980). Some of the most important works in the literature have been collected in Barnett and Serletis (2000) and Barnett and Binner (2004).
- Reimers et al. (2002) found that Divisia aggregates for several countries in Europe have better out-sample-predicting power for the GDP deflator in the Euro area. Schunk (2001) showed that using Divisia aggregates improves the accuracy of US real GDP and GDP deflator predictions.
- Binner (2005) finds there are strong indications that Divisia outperforms simple-sum aggregates in a non-linear framework when forecasting inflation for the euro.

- Methodology: Generalities
- This study compares direction of change forecasts produced by the GMLVQ algorithm. The first forecasts are produced using a set of macroeconomic variables (see below). Divisia monetary aggregates then replace simple-sum aggregates and the User Cost Price replaces short-term interest rates. Their performances are compared to forecasts produced by a benchmark model.
- What follows is a more detailed explanation of the methodology.

- Methodology: Divisia
- We know that the capital stock of money in a given time period is not equal to the monetary service flow (as capital goods do not fully depreciate in a period).
- The price of these monetary service flows is the opportunity cost, or user cost, of holding a particular monetary asset for that period.
- The User Cost Price then is the present value of however much interest an agent is not receiving because they are holding an asset, given that there exists a pure investment asset which provides a higher return and no monetary services.

- Methodology: Divisia
- The User Cost Price is calculated thusly:

$$\pi_{it} = (R_t - \gamma_{it})/(1 + R_t) \tag{1}$$

- where γ_{it} is the return on asset i and R_t is the return on the pure investment, or benchmark, asset.

- Methodology: Divisia
- With the User Cost Price precisely defined, an aggregate for the monetary service flows can be elaborated which will track these flows correctly. For this purpose a Divisia index is used. For the construction of Divisia indexes, let:

$$s_{it} = \pi_{it} m_{it} / \sum \pi_{jt} m_{jt} \tag{2}$$

- where m_{it} is the nominal monetary asset i at time t.

- Methodology: Divisia
- The Divisia monetary index is:

$$\ln M_t - \ln M_{t-1} = \sum_{t=1}^{n} s_{it}(\ln m_{it} - \ln m_{it-1}) \qquad (3)$$

- Here M_t is the quantity index and s_{it} is defined as $s_{it} = 1/2(s_{it} + s_{it-1})$. From the above equation, one can see that the growth rate of the index is a weighted sum of each monetary asset i. Each i has a share in the User Cost and this is precisely its corresponding weight in the Divisia index.

- Methodology: Divisia
- Finally, the accompanying User Cost Price index Π is defined as:

$$\ln \Pi_t - \ln \Pi_{t-1} = \sum_{t=1}^{n} s_{it}(\ln \pi_{it} - \ln \pi_{it-1}) \qquad (4)$$

- The idea here is that agents substitute toward holding the monetary assets which have the lowest relative user costs whenever there is a change in the own interest rate of another component monetary asset. This reflects how agents take into account opportunity costs in their decision process.

- Data Selection
- To assess the importance of monetary aggregation methods in regards to the exchange rates, it is essential to identify a set of variables that can be used as determinants during forecasting.
- This variables set must include the different monetary aggregates and component rates of interest.

- Following from Molinas Sosa, Binner and Tong (2021) these potential predictors were considered to be:
  1. Money supply (domestic and foreign),
  2. Interest rate (domestic and foreign),
  3. Output (domestic and foreign),
  4. Price levels (domestic and foreign),
  5. Current account balances (domestic and foreign),
  6. Oil price.

- LVQ
- Learning Vector Quantization (LVQ) constitutes a family widely used classification or ordinal regression algorithms, especially suitable for situations where model transparency and ease of interpretability is essential.
- The classifiers are parameterized by a set of prototypical-vectors which represent the classes in the input space. In the working phase, an unknown sample is assigned to the class represented by the closest prototype. Kohonen introduced the original LVQ scheme in 1986 (Kohonen, 1986).

- Assume training data $(x_i, y_i) \in \mathbb{R}^D \times \{1, ..., C\}$, $i = 1, 2, ..., N$, are given, D denoting the data dimensionality (number of input features) and C is number of different classes. An LVQ network consists of a number of prototypes $w_i \in \mathbb{R}^D$, $i = 1, 2, 3, ..., L$, living in the data space, each endowed with a class label $c(w_i) \in \{1, ..., C\}$. Note that $L = C$, if only one prototype per class is used. Given a distance measure (metric) $d(x, w)$ in $\mathbb{R}^D$, classification is based on a winner-takes-all scheme: a data point x is assigned the label $c(w_i)$ of the prototype closest (in metric $d(\cdot, \cdot)$) to x: $d(x, w_i) < d(x, w_j), \forall j \neq i$. Many LVQ variants use the usual squared Euclidean distance (w, c are column vectors),

$$d(x, w) = \| x - w \|_2^2 = (x - w)^T (x - w). \qquad (5)$$

- GLVQ
- In this contribution, we use modifications of Generalized LVQ algorithm (GLVQ, see Sato (1996)). In GLVQ, training is derived as a minimization of the cost function

$$E = \sum_{i=1}^{N} \Phi \left(\frac{d_{+}(x_i) - d_{-}(x_i)}{d_{+}(x_i) + d_{-}(x_i)} \right) \tag{6}$$

based on the steepest descent method. Φ is a monotonic function.

- However, GLVQ suffers from the problem that classification is based on a predefined metric.
- The algorithm uses the standard squared Euclidean distance which assumes all that input features are equally important when they may not be
- Only a linear subspace of the input space $\mathbb{R}^D$ may be needed to successfully perform the classification.

- GMLVQ
- Generalized Matrix LVQ is an extension of GLVQ which includes metric learning. The algorithm learns a generalized form of the squared Euclidean distance obtained by introducing a metric tensor $\mathbf{\Lambda}$ of adaptive weight factors.

- GMLVQ
- Generalized Matrix LVQ (GMLVQ) is an extension of GLVQ which includes metric learning. The algorithm learns a generalized form of the squared Euclidean distance,

$$d_{\mathbf{\Lambda}}(x, w) = (x - w)^T \mathbf{\Lambda}(x - w), \tag{7}$$

obtained by introducing a metric tensor $\mathbf{\Lambda}$ of adaptive weight factors. To guarantee that $d_{\mathbf{\Lambda}}$ is a metric, the $(D \times D)$ matrix $\mathbf{\Lambda}$ needs to be positive-definite. This can be achieved by substituting $\mathbf{\Lambda} = \mathbf{\Omega}^T\mathbf{\Omega}$, where $\mathbf{\Omega} \in \mathbb{R}^{D\times D}$, full rank. Note that the classification decisions will stay the same if the metric tensor is multiplied by any positive constant. Hence, the classifier is invariant under positive scaling of $\mathbf{\Omega}$. To remove this degeneracy, $\mathbf{\Lambda}$ is normalized to unit trace after each learning step ($\sum_i \mathbf{\Lambda}_{ii} = 1$).

- GMLVQ extends the cost function E by this more general metric and adapts the metric parameters $\mathbf{\Omega}_{ij}$ together with the prototypes by means of a stochastic gradient descent.
- Crucially, the learnt metric tensor quantifies feature relevances for the classification task through diagonal elements of $\mathbf{\Lambda}$, as well as provides information about the "optimal" input subspace sufficient for the classification.

- Methodology: Out-of-sample performance evaluation
- Model performances were evaluated against a simple baseline always predicting the majority class. This would be optimal strategy if the time series of test class labels was generated i.i.d. from a memoryless stationary source.
- For example, if in the training part we see more positive moves (class $+1$) then negative ones (class -1), then disregarding the classification input, the baseline would always predict class label $+1$.

- It is instructive to report both the out-of-sample classification accuracy of the models and the percent improvement over the baseline:

$$\frac{A_M - A_B}{A_B} \times 100\%, \qquad (8)$$

where A_M and A_B are the model and baseline accuracies.

Exchange Rate Direction of Change Forecasting and Divisia Monetary Aggregates

- Since the exchange rate data cannot be assumed to be normally distributed, the non-parametric paired Wilcoxon signed-rank test was required. Imperatively, each rolling forecast across comparable model variations needed to be performed on the same splits to allow for this paired test.
- The idea is to see if the difference between forecasts are statistically significant or simply due to chance.

- Data
- For this paper, the data are monthly series of the different variables in the models starting in January, 2001 through January, 2021 for all exchange rates, except those including CNY which end in 2018 (due to data availability).
- Data for Divisia aggregates are only available starting in January, 2001 at the Bruegel Institute and so the series begins at that particular date.
- The in-sample period goes from that date until the first quarter of 2009 and the out-of-sample period starts in the second quarter of 2009.

- Data
- Most of the data were retrieved from the St. Louis Fed Federal Reserve Economic Data (FRED), except:
- The simple-sum M3 monetary aggregate comes from the Organization for Economic Co-operation and Development's (OECD) database. The User Cost Price and the Divisia M3 Monetary Aggregates for the US were taken from the Center for Financial Stability's website. Divisia monetary aggregates and User Cost Prices for China were obtained from the University of Wuhan. Divisia monetary aggregates and User Cost Prices for Mexico were kindly given to us by...

- Data (continued)
- The Euro simple-sum M3 monetary aggregate also comes from the OECD database.
- The Current Account Balance was taken from the European Central Bank (ECB) database. The User Cost Price and Euro Divisia M3 monetary aggregates were taken from the Bruegel Institute database.

- Results
- In an effort to provide an extensive evaluation of the effectiveness of Divisia monetary aggregates in exchange rate forecasting, there were three separate experiments performed using the outlined modelling process.
- Simple Sum vs Divisia Datasets: The first consisted of every model variation performed on two separate datasets; one with the simple sum components and the reference interest rate alongside output, CPI, current account balances and oil prices; and then another with the simple sum monetary aggregates replaced by the Divisia index and the reference interest rate replaced by the User Cost Price.

- Results
- Simple Sum vs Divisia Variables: The second required every model to be estimated twice; at first with only the simple sum monetary aggregates and interest rates, and a second time with these replaced by the Divisia index and the User Cost Price, respectively. The goal of this experiment is to compare the specific performance of these the two pairs of variables when set against one another.
- Simple Sum and Divisia Variables: The last was comprised of all model variations trained merely on a combination of simple sum and Divisia money alongside interest rates and the User Cost Price. The main objective here is to obtain a classification of the importance the variables (how often the model "picks out" each variable)

- Results Summary

Figure: Number of times each model outperformed the other with statistical significance at 10*, 5** and 1*** percent in parenthesis.

Exchange rate	Experiment 1		
	Divisia	Simple-sum	Tie
EUR/USD	10(3**)	8(0)	0
MXN/EUR	5(0)	10(0)	3
CNY/EUR	8(0)	3(1*)	7
EUR/CHF	11(1*)	5(0)	2
USD/MXN	8(1*)	6(0)	4
USD/CNY	9(3***)	1(0)	8
USD/CHF	10(0)	8(0)	0

- Results Summary

Figure: Number of times each model outperformed the other with statistical significance at 10*, 5** and 1*** percent in parenthesis.

Exchange rate	Experiment 2		
	Divisia	Simple-sum	Tie
EUR/USD	14(1**, 2***)	4(0)	0
MXN/EUR	6(1*)	8(2*)	4
CNY/EUR	13(1*)	2(0)	3
EUR/CHF	14(1**, 3***)	1(0)	3
USD/MXN	10(2*)	5(1*)	3
USD/CNY	13(1***)	2(0)	3
USD/CHF	12(2*, 2**)	4(0)	2

- Results: USD/EUR

Figure: Classification accuracy (%) of the baseline model

Horizon	Training Size		
	18	24	30
1	49.102	52.096	49.701
3	58.084	58.683	48.503
6	61.677	53.892	47.904
12	59.880	55.689	51.497
24	55.689	63.473	63.473
36	77.246	73.653	71.257

- Results: USD/EUR

Figure: Classification accuracy (%) of the GMLVQ model trained on the simple sum dataset (SS-D) alongside the equivalent trained on the Divisia dataset (D-D)

Horizon	Training Size					
	18		24		30	
	SS-D	D-D	SS-D	D-D	SS-D	D-D
1	64.072	61.677	59.880	64.671	62.874	62.275
3	77.844	78.443	74.850	80.240	76.647	75.449
6	88.024	85.629	83.234	84.431	84.431	85.629
12	89.222	89.820	88.024	88.623	87.425	88.623
24	90.419	89.820	86.826	91.018	85.030	90.419
36	88.623	87.425	89.222	87.425	89.222	88.623

Figure: (%) improvement SS-D vs D-D compared to benchmark, w/ p-values

Horizon	Training Size					
	18		24		30	
	SS-D	D-D	SS-D	D-D	SS-D	D-D
1	30.488 0.006	25.610 0.018	14.943 0.163	24.138 0.018	26.506 0.014	25.301 0.023
3	34.021 0.000	35.052 0.000	27.551 0.001	36.735 0.000	58.025 0.000	55.556 0.000
6	42.718 0.000	38.835 0.000	54.444 0.000	56.667 0.000	76.250 0.000	78.750 0.000
12	49.000 0.000	50.000 0.000	58.065 0.000	59.140 0.000	69.767 0.000	72.093 0.000
24	62.366 0.000	61.290 0.000	36.792 0.000	43.396 0.000	33.962 0.000	42.453 0.000
36	14.729 0.001	13.178 0.008	21.138 0.000	18.699 0.001	25.210 0.000	24.370 0.000

Figure: Improvement accuracy (%) of Divisia dataset vs. simple sum dataset w/ p-values of the statistical test

Horizon	Training Size 18	Training Size 24	Training Size 30
1	−3.738 0.516	8.000 0.217	−0.952 0.873
3	0.769 0.782	7.200 0.020	−1.563 0.683
6	−2.721 0.102	1.439 0.637	1.418 0.564
12	0.671 0.655	0.680 0.705	1.370 0.527
24	−0.662 0.655	4.828 0.020	6.338 0.013
36	−1.351 0.480	−2.013 0.180	−0.671 0.705

- Results: Feature relevance
- In the first experiment, the User Cost Price was usually the third most relevant variable (feature) and tended to be picked out more than the other variables.
- More importantly, when compared against one another in the third experiment, the User Cost Price was always the more relevant variable, above the reference rate, the Divisia aggregate and the simple-sum aggregate (in that order).

- Discussion
- What these results would seem to indicate is that, on the one hand, agents are not just monitoring short-term interest rates but the returns available on a variety of assets in different time periods; on the other hand, agents take into account the opportunity cost of holding assets - the foregone return.
- These results are also consistent with previous work showing that interest rates are good exchange rate predictors.
- The most important predictors of exchange rate movements appear to be GDP and BOP.

Conclusions

- Forecasting exchange rates is not a lost cause.
- Interest rates seem to have become one of the main driving variable for exchange rates (along GDP and BOP).
- Machine learning methods such as GMLVQ are getting researchers closer to the holy grail.
- The use of Divisia monetary aggregates and the User Cost Price can help improve forecasts.

Exchange Rate Direction of Change Forecasting and Divisia Monetary Aggregates

Jane Binner, Luis Molinas Sosa, Megan Hart, Peter Tino

Department of Financial Inclusion
Central Bank of Paraguay

September, 2023