# Economic and political effects on currency clustering dynamics

Marcel Kremer*†[a,b,c], Alexander P. Becker*‡[c,d], Irena Vodenska§[d], H. Eugene Stanley¶[c] and Rudi Schäfer‖[b,e]

[a] *Chair for Energy Trading and Finance, University of Duisburg-Essen, Universitätsstraße 12, 45141 Essen, Germany*
[b] *Faculty of Physics, University of Duisburg-Essen, Lotharstraße 1, 47048 Duisburg, Germany*
[c] *Center for Polymer Studies and Department of Physics, Boston University, 590 Commonwealth Avenue, Boston, MA 02215, USA*
[d] *Department of Administrative Sciences, Metropolitan College, Boston University, 1010 Commonwealth Avenue, Boston, MA 02215, USA*
[e] *Research and Prototyping, Arago GmbH, Eschersheimer Landstraße 526-532, 60433 Frankfurt am Main, Germany*

October 15, 2018

Revised version published in *Quantitative Finance, 2018.*

### Abstract

We propose a new measure named the *symbolic performance* to better understand the structure of foreign exchange markets. Instead of considering currency pairs, we isolate a quantity that describes each currency's position in the market, independent of a base currency. We apply the $k$-means++ clustering algorithm to analyze how the roles of currencies change over time, from reference status or minimal appreciations and depreciations with respect to other currencies to large appreciations and depreciations. We show how different central bank interventions and economic and political developments, such as the cap on the Swiss franc to the euro enforced by the Swiss National Bank or the Brexit vote, affect the position of a currency in the global foreign exchange market.

**Keywords:** Symbolic performance; Currency clustering; Currency dynamics; Central bank intervention

**JEL Classification:** E52; E58; G15

*MK and APB share first authorship.
†marcel.kremer@uni-due.de (corresponding author)
‡apbecker@bu.edu (corresponding author)
§vodenska@bu.edu
¶hes@bu.edu
‖rschaefer@arago.de

# 1 Introduction

The structure of the foreign exchange market (FX market) poses a special challenge to traders and researchers: Whereas stocks or commodities are quoted in the domestic currency of their market, currencies are traded worldwide and in pairs, that is, in terms of a variety of other currencies. This implies that we cannot compare two currencies without defining beforehand the transaction currency. The choice of the transaction or base currency influences the currency research results in two important ways: One, the underlying fundamentals of the base currency are changing, providing a non-constant background. Two, the base currency introduces a complex interdependency through its relation with the two other currencies. As a consequence of this triangular structure, the FX market is a system in which the appreciation of one currency implies depreciation of another one. Likewise, movements in one currency pair can and will propagate into the time series of other currency pairs. Therefore, any traditional form of analysis hinges on the choice of the base currency.

In this paper, we present a novel approach to alleviate this issue. Instead of considering currency pairs, we treat each currency $i$ in a market of $K$ currencies as an individual entity with assigned symbolic performance $\zeta_i$. This approach introduces a measure independent of base currency to investigate the hierarchy and the dynamics of the FX market.

We aim to encode the relationship of each currency with all other $K-1$ currencies into one quantity. Instead of considering $K-1$ currency pairs for currency $i$, we compress information of these $K-1$ pairs into one quantity for each currency by measuring the relative performance of currency $i$ in relation to other currencies. As a result, the symbolic performance of a currency quantifies its global role within the entire foreign exchange market.

Using the symbolic performance, we investigate how a currency's role evolves within the market in the wake of changing economic conditions. As exchange rates are affected by monetary policy, we especially consider central banks' currency interventions that may be conducted directly, for example, if a central bank purchases or sells the domestic currency. In more extreme cases central banks may introduce a cap or a peg of its currency, backing this policy by currency transactions. Data for interventions of this kind are publicly available for the Swiss franc, the Mexican peso, the Singapore dollar and Japanese yen among the currencies considered in this paper.

In the literature, the effects of central bank interventions (CBIs) on foreign exchange rates have been studied by various techniques, particularly focusing on volatility of exchange rates. These techniques include GARCH type approaches (Almekinders and Eijffinger, 1996; Baillie and Osterberg, 1997b,a; Dominguez, 1998; Beine et al., 2002), implied volatility estimation of currency options (Bonser-Neal and Tanner, 1996; Dominguez, 1998), regime-switching analysis of mean and variance of exchange rates (Beine et al., 2003), realized volatility estimation (Dominguez, 2006; Beine et al., 2009; Cheng et al., 2013), time series study of news reports (Fatum and Hutchison, 2002), and event study of CBIs (Fatum and Hutchison, 2002, 2003; Fatum, 2008).

Most of the works quoted consider only three currencies – the German deutschmark (euro), Japanese yen and US dollar – in line with the respective CBI data of the German Bundesbank (European Central Bank), Bank of Japan and Federal Reserve System. Our approach, however, links currencies to one another and thus explicitly incorporates exchange rate data of currencies whose central banks did not intervene in the time period being analyzed. This methodological distinction allows us to examine not only effects of CBIs on the domestic currencies, but on the currencies embedded in the FX market.

The rest of this paper is structured as follows. We present our foreign exchange data involving 14 currencies in Section 2. We lay out the framework and the methodology of obtaining the symbolic performance measure in Section 3. In Section 4, we study the statistics of symbolic performances for the entire time horizon as well as for specific subintervals. In particular, we present the results of our cluster analysis revealing the temporal evolution of the symbolic performances and identifying different roles currencies play within the FX market. We link changes in roles and behaviors of currencies to central bank interventions as well as economic shocks. We offer our discussion and conclusion in Section 5.

# 2 Data

We use python to download currency exchange rates from OANDA, which provides an open access API [1]. Our data comprises the following currencies: Australian dollar (AUD), Canadian dollar (CAD), Swiss franc (CHF), British pound (GBP), Hong Kong dollar (HKD), Japanese yen (JPY), Mexican peso (MXN), Norwegian krone (NOK), New Zealand dollar (NZD), Swedish krona (SEK), Singapore dollar (SGD), US dollar (USD), and South African rand (ZAR). All currencies are quoted in terms of euro (EUR). This gives us 14 distinct exchange rate time series when we include a dummy euro time series which consists of only ones. The data set spans from January 2nd, 2005, to May 9th, 2017, a period of more than twelve years. The start date is chosen such that we can observe a time window as long as possible while at the same time maintaining the quality of the data. Exchange rates are quoted in 10-minute intervals. There is no trading on certain holidays as well as from Friday night to Sunday night.

We treat missing data as follows: If no euro pair has been traded at all in a given 10-minute interval, the OANDA platform does not report any exchange rates for this time point. We exclude these time points from our analysis. If one particular pair is not traded in a given 10-minute interval, the OANDA platform records a return of zero. If more than two euro pairs have not been traded, we discard this time point. If at most two euro pairs have not been traded, we impute data by drawing values from a normal distribution. We estimate the mean and standard deviation of that normal distribution as average and standard deviation of empirical log returns of the missing currency in the past 24 hours. If data are missing and we cannot estimate average and standard deviation over the last 24 hours, we exclude the time point from our analysis.

The data set contains 444,360 intervals, and after we account for missing data according to the above described procedure, approximately 413,000 intervals per currency remain. Since we analyze 14 currencies, this yields about 5.8 million exchange rates which we use for our study.

We gather information on central bank interventions from the websites of central banks who publish their activity as well as news reports from sources like Reuters[2].

# 3 Methodology

## 3.1 Exchange rates and returns

We study a financial market consisting of $K$ assets. Since we consider the foreign exchange market, each asset $i$ is a currency and will be linked to the remaining $K-1$ currencies through their exchange rate $S_{i,j}(t)$ at time $t$, where $j = 1, \ldots, K$, $j \neq i$. Obviously $S_{i,i} \equiv 1$. In matrix form, this becomes

$$\mathbf{S}(t) = \begin{pmatrix} 1 & S_{1,2}(t) & \cdots & S_{1,K}(t) \\ S_{2,1}(t) & 1 & \cdots & S_{2,K}(t) \\ \vdots & \vdots & \ddots & \vdots \\ S_{K,1}(t) & S_{K,2}(t) & \cdots & 1 \end{pmatrix}. \quad (1)$$

[1]http://developer.oanda.com

[2]Examples: http://www.banxico.org.mx/sistema-financiero/estadisticas/mercado-cambiario/banco-mexico-s-foreign-exchan.html and http://www.reuters.com/article/global-forex-boj-idUSL8N15Q4BY

Each row of this matrix describes exchange rates given a base currency, while each column of the matrix describes exchange rates given a counter currency. In the absence of a bid-ask spread, the reciprocal relationship implies $S_{i,j}(t) = 1/S_{j,i}(t)$.

Using the structure of the FX market, we can construct this matrix from the exchange rates of just one currency with all other currencies at time $t$. Under the assumption of no arbitrage, the exchange rates $S_{i,j}(t)$ and $S_{i,k}(t)$ imply the exchange rate $S_{j,k}(t)$ via $S_{j,k} = S_{j,i}(t)\, S_{i,k}(t)$. It then suffices to consider the exchange rate vector $\mathbf{S}_i(t)$ at time $t$ which comprises of the exchange rates of currency $i$, acting as a base currency, with all other currencies:

$$\mathbf{S}_i(t) = \left(S_{i,1}(t), S_{i,2}(t), \ldots, S_{i,K}(t)\right). \tag{2}$$

It carries complete information of the system, in that the outer product of the vector and its inverse yields the exchange rate matrix in Eq. (1),

$$\mathbf{S}(t) = \begin{pmatrix} S_{1,i}(t) \\ S_{2,i}(t) \\ \vdots \\ S_{K,i}(t) \end{pmatrix} \cdot \left(S_{i,1}(t), S_{i,2}(t), \ldots, S_{i,K}(t)\right). \tag{3}$$

In this study, we calculate logarithmic returns between consecutive exchange rates in a time interval $\Delta t$:

$$R_{i,j}(t) = \log S_{i,j}(t) - \log S_{i,j}(t - \Delta t). \tag{4}$$

Analogous to Eq. (1), we can write the exchange rate returns in matrix form:

$$\mathbf{R}(t) = \begin{pmatrix} 0 & R_{1,2}(t) & \cdots & R_{1,K}(t) \\ R_{2,1}(t) & 0 & \cdots & R_{2,K}(t) \\ \vdots & \vdots & \ddots & \vdots \\ R_{K,1}(t) & R_{K,2}(t) & \cdots & 0 \end{pmatrix}. \tag{5}$$

Here the reciprocal relationship becomes $R_{i,j}(t) = -R_{j,i}(t)$.

## 3.2 Symbolic performance

Considering the foreign exchange return matrix in Equation (5), it is intuitively clear that a base currency with a large number of positive returns appreciates overall, whereas a base currency with a large number of negative returns depreciates overall. This information for a given currency $i$ is stored in the $i$th row of the return matrix.

We introduce the symbolic performance $\zeta_i(t)$ which we define as the difference between the number of positive and negative returns currency $i$, acting as base currency, has at time $t$. This is achieved by applying the sign-function to returns $R_{i,j}(t)$ in the return matrix, yielding $+1$ for positive returns and $-1$ for negative returns, and then summing row-wise:

$$\zeta_i(t) = \sum_{j=1}^{K} \operatorname{sgn} R_{i,j}(t). \tag{6}$$

We know that $\operatorname{sgn} R_{i,j}(t) = -\operatorname{sgn} R_{j,i}(t)$. In other words, if the currency $i$ rises with respect to the currency $j$, then currency $j$ falls with respect to currency $i$.

The symbolic performance vector then lists the symbolic performances $\zeta_i(t)$ of all currencies at time $t$. In matrix notation,

$$\boldsymbol{\zeta}(t) = \operatorname{sgn} \mathbf{R}(t) \cdot \mathbf{1} = (\zeta_1(t), \zeta_2(t), \dots, \zeta_K(t)) \,. \tag{7}$$

By construction, the symbolic performance can take values from $-K+1$ to $K-1$ in steps of 2, and for a given time $t$ each value appears exactly once; this means $\sum_i \zeta_i(t) = 0$. As a consequence, if currency $j$ has a larger return than currency $k$ with respect to the base currency $i$, then currency $j$ will appreciate with respect to currency $k$.

It is important to point out that the symbolic performance is a measure for each asset itself: We get $K$ values $\zeta_i(t)$, $i = 1, \dots, K$, which describe the behavior of the $K$ currencies individually. While they contain the correlation structure of the FX market, the symbolic performances are otherwise independent, except that $\zeta_i(t) = \zeta_j(t)$ if and only if $i = j$.

In a simple model, we can decompose the variance of a given return time series into one part which corresponds to the state of the market, i.e., the sum of the variance of all returns, and into another part which corresponds the idiosyncratic variance. While the symbolic performance of a given currency contains information about all other currencies, it does not provide information about the magnitude of an average return at a given time. In other words, it disregards the variance of the return time series due to the state of the market. However, it retains information on the idiosyncratic variance of the return time series of currencies. It is easy to see that Eq. (6) is equivalent to a ranking of returns given a fixed base currency $i$ where the largest positive (negative) return is assigned the largest positive (negative) value for $\zeta_i(t)$:

$$\begin{aligned} \zeta_i(t) &= \sum_{j=1}^{K} \operatorname{sgn} R_{i,j}(t) \\ &= \sum_{j=1}^{K} \left[2\,\mathrm{H}\left(R_{i,j}(t)\right) - 1\right] = 2 \sum_{j=1}^{K} \mathbb{1}\left(R_{j,1}(t) \leq R_{i,1}(t)\right) - 1 - K \\ &= 2 \operatorname{rank} R_{i,1}(t) - (K+1), \end{aligned} \tag{8}$$

where H denotes the Heaviside step function and $\mathbb{1}$ the indicator function. Note that this is true regardless of our choice of base currency $i$, as pointed out earlier.

Currencies with larger variance relative to the other currencies are more likely to exhibit large swings and thus higher magnitudes of $\zeta$, and vice versa. As a result, currencies that tend to be more volatile regardless of market state tend to have more fat-tailed symbolic performance distributions $P(\zeta_i)$. Currencies that take a position at the center of the market in comparison to the remaining currencies tend to have a symbolic performance distribution $P(\zeta_i)$ that is reminiscent of a Gaussian. For the purpose of finding the position of a currency, it is then sufficient to consider the distribution of the absolute values of the symbolic performance, $P(|\zeta_i|)$, a discrete probability distribution which describes how often a currency takes the symbolic performance $-(K-1)$ or $K-1$, $-(K-3)$ or $K-3$, etc.

### 3.3 $k$-means++ clustering

The movement of currencies is closely linked to the macroeconomic developments in the corresponding countries. Therefore the more qualitative description of currencies as, for example, G4 currencies, G10 currencies or commodity currencies, is valid on long time scales. While this classification helps describe the roles and behaviors of currencies in general, economic shifts as well as political or monetary shocks can alter the position of a currency within the market on varying time scales, and the symbolic performance is able to capture these shifts. To investigate these currency dynamics, we evaluate the symbolic performance distributions $P_t(\zeta_i)$ on reasonably short time scales, from $t$ to $t + \Delta T$, and classify them according to different currency behavior.

We use $k$-means++ clustering for this task. Cluster analyses find application in many fields, such as, biology and bioinformatics, business and marketing, medicine, social networks, computer science, climate and weather research, and data mining, where they are used to discover similarities and patterns in large amounts of data. The $k$-means++ clustering algorithm takes as input a set of vectors as well as a number of clusters or classifications $k$. It then determines $k$ points which serve as cluster centroids such that the overall distance between all data points and their respective cluster center is minimized. This algorithm itself does not require any knowledge of the data set to operate, and it will find a solution even if the data is not properly clustered into $k$ centers. Therefore, the choice of the number of clusters $k$ becomes critical, as we will explain later in this section.

The algorithm schematically works as follows: Given $N$ different points in $d$-dimensional space, $k$ random points are selected as the initial cluster centers $\mu_i$, $i = 1, \ldots, k$. Each of the $N - k$ remaining points are then associated with the cluster center to which they are closest. After assigning all points to clusters, the cluster centers, defined as the mean of all points in the cluster, will change. The cluster centers are thus re-evaluated before the next iteration, and all points are re-classified to be closest to the new cluster centers, again. This procedure is reiterated until convergence is reached, that is, all points $\mathbf{x}$ are distributed in clusters $C_i$, $i = 1, \ldots, k$, with centers $\mu_i$ such that

$$\sum_{i=1}^{k} \sum_{\mathbf{x} \in C_i} ||\mathbf{x} - \mu_i||^2 \tag{9}$$

is minimized.

The $k$-means++ algorithm uses the squared Euclidean metric as its distance measure. Recall that we are interested in the classification of symbolic performance distributions $P_t(\zeta_i)$, estimated on time intervals $[t, t+\Delta T]$. To this end we arrange the values of this discrete probability distribution $P_t(\zeta_i)$, i.e., the relative frequencies $\mathrm{Prob}_t(\zeta_i)$, in a vector,

$$\mathbf{P}_t(|\zeta_i|) = \begin{pmatrix} \mathrm{Prob}_t(\zeta_i & = & \pm(K-1)) \\ \mathrm{Prob}_t(\zeta_i & = & \pm(K-3)) \\ & \vdots & \\ \mathrm{Prob}_t(\zeta_i & = & \pm 1) \end{pmatrix}. \tag{10}$$

Each vector $\mathbf{P}_t(|\zeta_i|)$ is a composition. This has implications regarding the distance measure. For compositional data the use of Aitchison's distance, introduced in Aitchison (1982), is much more appropriate. In contrast, the use of Euclidean distance may be problematic, as Martín-Fernández et al. (1998) and Aitchison et al. (2000) point out. Alternatively, we can apply an isometric log-ratio (ilr) transformation to $\mathbf{P}_t(\zeta_i)$. The so-transformed data is then accessible to analytical methods that use the Euclidean metric. Since the scikit-learn package which we use to classify our data requires the Euclidean distance as the distance metric for $k$-means++, we apply an ilr transformation and cluster the results. In the last step, we perform the inverse ilr transformation to express the cluster centers in terms of relative frequencies again[3].

In order to be able to provide a meaningful interpretation of the symbolic performance distribution, $\Delta T$ needs to be carefully chosen considering three criteria. One, there is no trading on the OANDA platform for a few hours every weekend. In order to avoid capturing beginning- or end-of-the-week effects, it is sensible that each time window is one week long or multiples of one week. Two, if we choose the length of the window to be too small, we undersample the

[3] By using the absolute values of the symbolic performances, we reduce the dimensionality from $K$ to $K/2$, and by applying the ilr transformation, we further reduce it to $K/2-1$. This mitigates the problem of sparsity of high-dimensional spaces significantly.

distribution. Three, if $\Delta T$ is too large, we may miss local trends and fluctuations, and we may limit our ability to resolve the effects of political or economic shocks whose impact can last anywhere from a few days to several months or years. Given our data resolution of 10 minutes, $\Delta T = 2$ weeks appears to be the appropriate choice to satisfy our criteria.

It is worth pointing out that we perform the cluster analysis on the pool of all symbolic performance distributions of all currencies, that is, on the entire FX market. As a result, information about each currency as well as each point in time informs the classification process. In other words, if two different currencies are classified to belong to the same cluster at different times, their respective behavior or role in the market is very comparable.

As pointed out previously, we also have to decide on a reasonable number of clusters $k$, based on the data set. We want to choose as few clusters as possible with as large explanatory power as possible. The gap statistic, introduced by Tibshirani et al. (2001), is a useful metric to determine the appropriate value for $k$ as suggested by the structure of the data. Since $k$-means++ does not penalize model complexity, that is, the number of clusters to be found, increasing $k$ does not increase the value of the cost function. This is true whether the data is clearly clustered or not. The gap statistic contrasts the benefit of adding one more cluster to a structured data set to the benefit of adding one more cluster to a comparable but random and unclustered data set. The gap is the difference in the cost function of clustering these two data sets. We increase $k$, starting with just one cluster, until the size of the gap reaches its first local maximum; at this point adding the $k + 1^{\text{th}}$ cluster provides diminishing benefits. After performing this analysis on our data set, we conclude that separating our data in six clusters is the most sensible choice.

# 4 Empirical results

## 4.1 Overall symbolic performance distributions

Based on the foreign exchange data introduced in Section 2 and the symbolic performances calculated according to Eq. (6), we analyze the overall symbolic performance distributions $P(\zeta_i)$, of all currencies $i = 1, \ldots, 14$, as shown in Figure 1. These discrete probability functions indicate how often each currency $i$ has taken each value $\zeta_i$ from $-13$ to $+13$ during the entire time period from 2005 to 2017.

We observe that the distributions exhibit characteristic shapes for different currencies. Some are clearly convex, like the distributions for the Japanese yen, the New Zealand dollar, and the South African rand. Others are clearly concave, like the distributions for the euro, the Hong Kong dollar, the Singapore dollar, and the US dollar. This reflects the degree to which extent currencies tend to exhibit large swings, that is, have large volatility relative to the remaining currencies regardless of market conditions on one hand, or to which currencies tend to stay in the center of the market on the other. Irrespective of their shapes, however, all distributions are quite symmetric around 0[4].

We describe the distributions based on the appearance of maxima in the center or at the tails; in other words we can distinguish the currencies' performance distributions by their curvature:

1. Strongly concave: EUR, HKD, SGD, USD.
2. Slightly concave: CHF, GBP.
3. Fairly flat: CAD, SEK.
4. Slightly convex: AUD, MXN, NOK.
5. Strongly convex: JPY, NZD, ZAR.

[4] We make use of this feature in our cluster analysis by considering only the absolute value of the symbolic performance and its distribution.

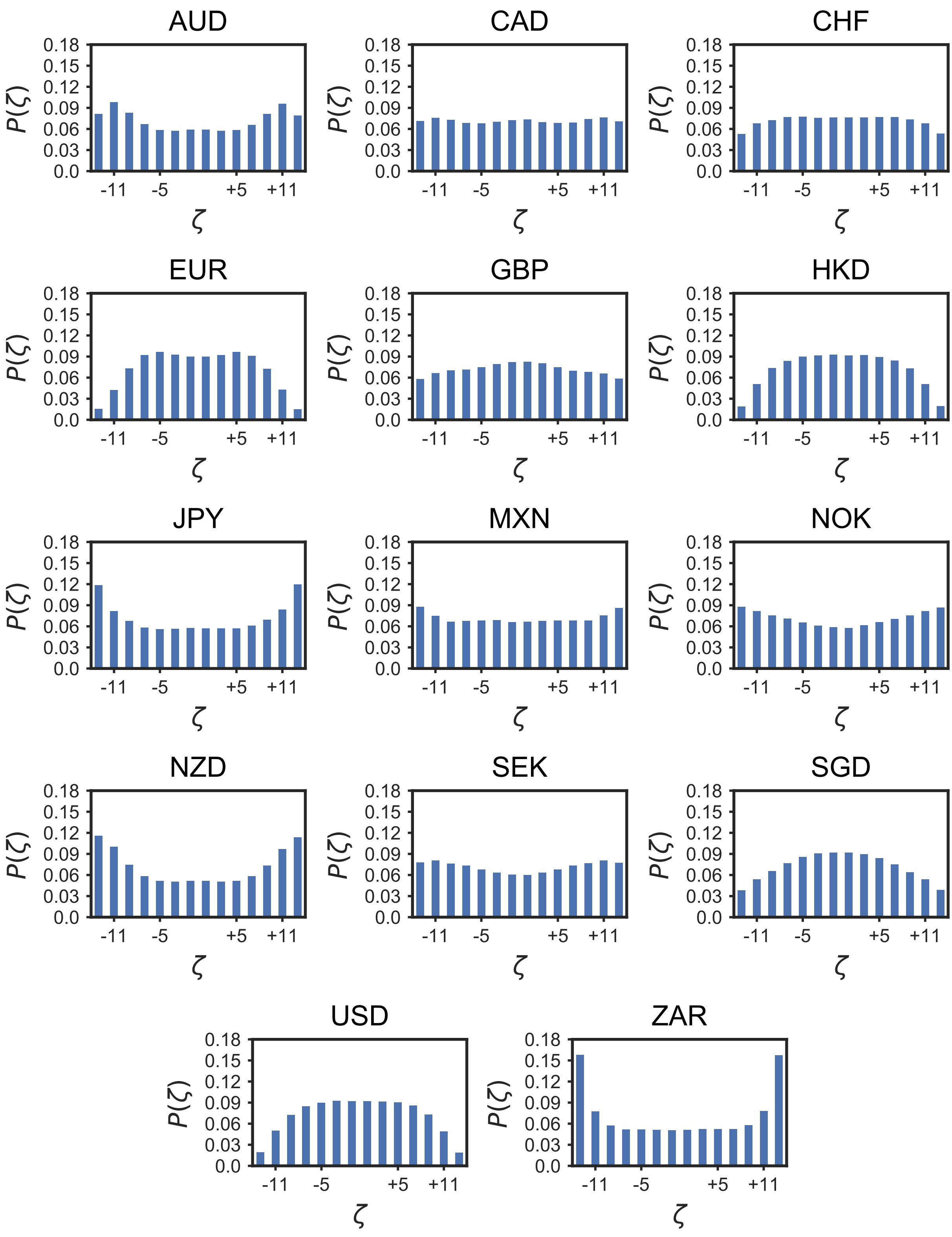


Figure 1: Overall symbolic performance distributions $P(\zeta_i)$, $i = 1, \ldots, 14$, for the currencies in our data set. We can distinguish between a few different shapes, characterized by their curvature: Some currencies exhibit high probabilities at the center and small probabilities in the tails. Other currencies' probability mass functions are fairly uniform. Some currencies exhibit comparably small probabilities at the center and high probabilities in the tails. At the extremes we find the euro (EUR) with very concave curvature and the South African rand (ZAR) with very convex curvature.

Let us first consider the tail-heavy distributions. Currencies like the ZAR with an overall convex symbolic performance distribution are most likely to exhibit returns that are larger in magnitude than all other returns. These currencies tend to outperform or underperform the rest of the market. This does not, however, imply that these currencies necessarily are currencies with large volatility. We can easily change the base pair to change the volatility of a currency. When measured in euro, the Swiss franc, for example, is the currency with the lowest volatility, but when measured in US dollar, the Swiss franc exhibits higher volatility. This underscores the usefulness of the symbolic performance as a measure of relative volatility since it takes into consideration all currency pairs at once.

Currencies like the EUR or USD with strongly concave symbolic performance distributions tend to maintain a position at the center of the market. This means that they are unlikely to consistently strongly appreciate or depreciate against other currencies. Instead, they serve as some form of reference against which other currencies are held. Consider an imaginary currency with true reference status. The size of its returns would be distributed like a normal distribution centered around zero with a very small standard deviation due to minor fluctuations. As a consequence, large fluctuations would be very unlikely, and likewise the chance for being at the extreme positions of the market would be very small as compared to other currencies.

We also identify currencies which appear to take all symbolic performance values and therefore roles in the market with similar frequency. This can be part of the behavior of the currencies, or it can be a mixture of periods in which their symbolic performance distributions exhibit concave curvature and periods in which they exhibit convex curvature.

### 4.2 Dynamics of symbolic performance distributions

According to the 2016 Triennial Central Bank Survey of FX and over-the-counter (OTC) derivatives markets, the daily trading volume of the foreign exchange market exceeds 5 trillion USD.[5] This enormous volume results in high liquidity and rich dynamics. Furthermore, political and economic interventions can result in shifts in the structure of the market. The clustering procedure introduced in Section 3.3 is able to reveal the described dynamics and shifts. The results of the cluster analysis are presented in Figs. 2 and 3. Figure 2 shows the cluster centers found by the $k$-means++ algorithm, that is, the symbolic performance distributions $P(\zeta)$ each cluster corresponds to, which allows us to interpret the classification of currencies. Figure 3 exhibits how the currencies evolve through the clusters over time with a resolution of $\Delta T = 2$ weeks.

Table 1 offers a summary of how often each currency is classified into each cluster center during the twelve year period. The euro and the US dollar occupy clusters 1 and 2 most often. The Japanese yen, the New Zealand dollar, and the South African rand can be found at the opposite end, as they are most often classified into clusters 5 and 6.

We can distinguish two different groups of currencies: Currencies that maintain their role in the foreign exchange market over a long time and spend the majority of the observation period in one or two clusters, and currencies that change their role significantly over time, either in the form of short bursts or long-term shifts. A typical example of the first group is the euro which maintains its reference role almost throughout the entire time period. Typical examples for the second group are the British pound with the sudden change in clusters after the Brexit vote and the Swiss franc with drastic changes around 2011 (pegging to the euro) and 2015 (unpegging from the euro).

We further investigate the currencies with particularly interesting behaviors, such as Canadian dollar, Swiss franc, euro, British pound, Japanese yen, Mexican peso, Singapore dollar and US dollar.

[5] Bank for International Settlements, http://www.bis.org/publ/rpfx16.htm

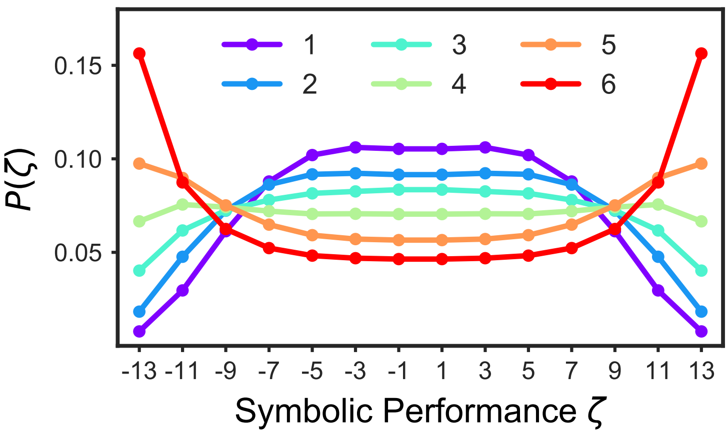


Figure 2: Symbolic performance distributions $P(\zeta)$ corresponding to the cluster centers found by the $k$-means++ algorithm. For the sake of better illustration, we draw a line for each discrete probability distribution, keeping in mind that $P(\zeta)$ is only defined for odd integers between $-13$ and 13. The violet distribution (cluster 1) corresponds to currencies which take central positions in the market, whereas the red distribution (cluster 6) corresponds to currencies which tend to take more extreme positions in the market.

| cluster | AUD | CAD | CHF | EUR | GBP | HKD | JPY | MXN | NOK | NZD | SEK | SGD | USD | ZAR |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0 | 0 | 1 | 49 | 0 | 24 | 0 | 0 | 0 | 0 | 0 | 0 | 27 | 0 |
| 2 | 0 | 0 | 10 | 39 | 1 | 53 | 0 | 1 | 0 | 0 | 0 | 24 | 50 | 0 |
| 3 | 0 | 30 | 49 | 9 | 53 | 22 | 4 | 10 | 4 | 0 | 8 | 65 | 21 | 0 |
| 4 | 32 | 35 | 27 | 3 | 37 | 1 | 23 | 47 | 30 | 3 | 42 | 10 | 3 | 0 |
| 5 | 65 | 33 | 12 | 0 | 7 | 0 | 29 | 34 | 64 | 60 | 49 | 0 | 0 | 16 |
| 6 | 3 | 2 | 2 | 0 | 1 | 0 | 44 | 9 | 2 | 37 | 1 | 0 | 0 | 84 |
| average | 4.7 | 4.1 | 3.5 | 1.7 | 3.5 | 2 | 5.1 | 4.4 | 4.6 | 5.3 | 4.4 | 2.9 | 2 | 5.8 |

Table 1: We report how frequently (in %) each currency appears in each cluster, enumerated from 1 to 6 as in Fig. 2, as well as the average cluster value. The euro has the lowest average cluster component, followed by the US dollar, the two main references in the FX market. The New Zealand dollar, the Japanese yen, and the South African rand stand out as the currencies with the most tail-heavy symbolic performance distributions.

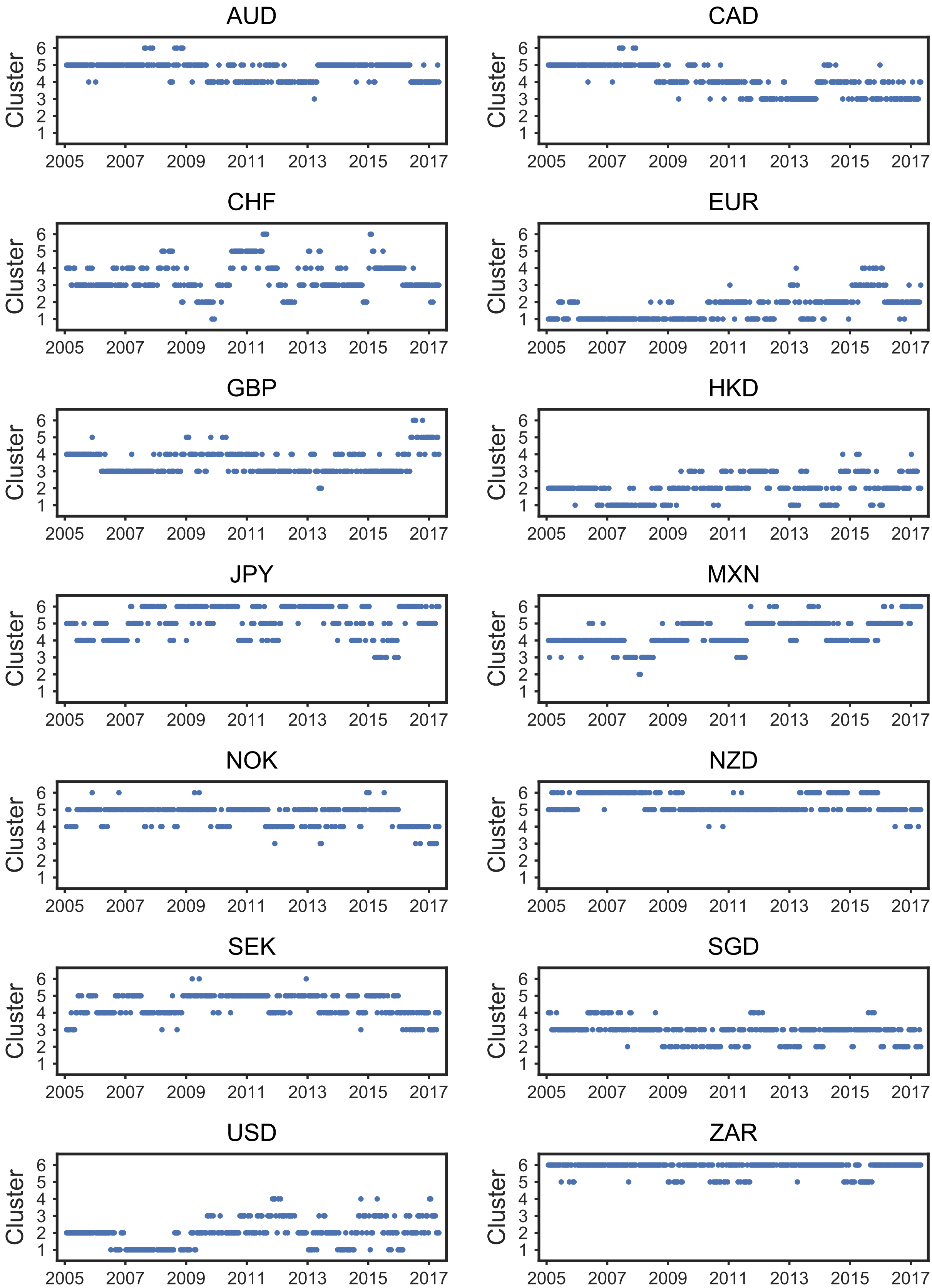


Figure 3: Currency dynamics according to the cluster association for the 14 currencies in our data set. Cluster 1 corresponds to reference-like behavior, whereas cluster 6 corresponds to heavy-tail behavior.

**CAD.** The Canadian dollar starts out as a currency that tends to stay at the tails of the market, as denoted by its long presence in cluster 5. Around late 2008, with the onset of the global financial crisis, the heavy tails become weaker, as shown by its presence in cluster 4. From mid 2011 until late 2013, the Canadian dollar behaves more like the Swiss franc or the British pound, as shown by its association to cluster 3. In 2014 this development appears to revert itself, briefly. Eventually, however, the Canadian dollar maintains clusters 3 and 4. It is crucial to emphasize that the symbolic performance distributions associated with the clusters provide relative information. Therefore, we cannot automatically infer whether the Canadian dollar itself has changed its role or the changes pertain to the rest of the market; instead we have to consider the development of other currencies during these time periods.

**CHF.** In Fig. 4 we zoom into the cluster association of the Swiss franc, drawing lines for each publicly announced central bank intervention of the Swiss National Bank (SNB). The Swiss franc exhibits particularly stable behavior prior to the onset of the global financial crisis in 2008, being in cluster 3 most of the time. This is consistent with its role as safe haven, and it is comparable to the behavior of the British pound and the Singapore dollar. With the onset of the financial crisis in 2008, the Swiss franc exhibits larger swings than other currencies for a while, but soon returns to cluster 3. Starting early 2009, the Swiss franc begins an extended period in cluster 2 which is usually only occupied by reference currencies or currencies which are in a managed floating regime. We can link this behavior to aggressive action of the SNB which acted to maintain the stability of the franc and curb its appreciation against other major currencies by buying euros and other currencies. These public interventions started in early 2009, and we consider them successful in that they moved the CHF into cluster 1, where we refer to currencies as reference currencies. In the wake of the European sovereign debt crisis, the SNB intervened more frequently, as indicated by the numerous lines in early 2010, as shown in Fig. 4. During the time of these interventions the Swiss franc stays in clusters 2 and 3 despite significant market pressure. However, eventually the SNB abandoned its policy to purchase euros and other currencies. In the following weeks, the impact that this policy had becomes obvious as the Swiss franc is mostly in cluster 5, corresponding to above-average volatility. A consequence of the end of central bank interventions was an overall appreciation of the Swiss franc in relation to the euro and the US dollar. These appreciations eventually shifted the Swiss franc to cluster 6 which it had never occupied before.

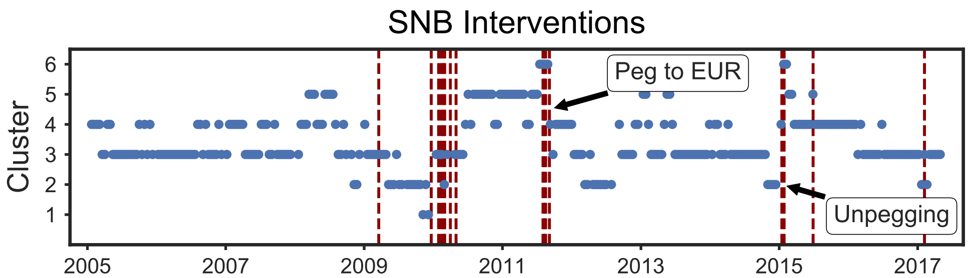


Figure 4: Cluster association of the Swiss franc, highlighting the time points of publicly announced central bank interventions.

These large currency movements sparked further central bank interventions, resulting in the enforcement of a cap on the exchange rate of the Swiss franc to the euro in September 2011 by the Swiss National Bank. Our analysis indicates that after a couple of months the market accepted the SNB's controlled cap, and the Swiss franc moved back to its more pre-crisis role, appearing in cluster 3 most of the time. Another intervention in January 2015, highlighted in Fig. 4, corresponds to the uncapping of the Swiss franc with respect to the euro. This SNB intervention pushed the Swiss franc in clusters 5 and 6 again. For roughly two months the Swiss franc exhibited heavy tails in the symbolic performance distribution. Over the time scale of a year, it returned to cluster 3, its standard position. Further intervention is suggested to have taken place in early February 2017, inferred from rising sight deposits[6], and we indeed observe a drop in cluster state around that period.

**EUR.** Our analysis confirms the euro as a major reference. This is detailed by its continued presence in cluster 1, despite global and local financial turmoil. During the European sovereign debt crisis, though, the euro leaves cluster 1 to move to cluster 2 more frequently. In July 2012, Mario Draghi announced that the ECB would do “whatever it takes” to protect the euro[7]. This led to a large appreciation of the euro in the following months. As a result of this, the euro leaves cluster 1 for almost a year, and it even briefly moves to cluster 3 and cluster 4, indicating a more volatile period, before returning to cluster 2 and then 1. During most of 2015, the euro is in cluster 3 again, with significant time in cluster 4. Concurrently, no currency belongs to cluster 1 for more than a few months at a time. We can interpret this as a lack of reference in the FX market. Even after the return to cluster 2 in early 2016, the euro occupies cluster 1 only for brief periods of time.

**GBP.** Being in cluster 3 for the majority of the time horizon, the British pound plays the role of a safe haven currency. On June 23, 2016, the British people voted to leave the European Union, causing a major shock to the world economy. This is reflected by a move of the British pound to cluster 4 already at announcement of the referendum date, as we observe in Fig. 5 in more detail, indicating higher volatility. In the weeks up to the referendum, uncertainty increased further and the pound was classified into cluster 5. The surprising outcome of the vote caused large movements in exchange rates around the globe, especially involving the British pound. With our methodology we can monitor if the British pound will return to its role or if this kind of political intervention changes the hierarchy in the foreign exchange market in the long run. It will be interesting to analyze whether the outcome of the British referendum to leave the European Union will have a profound and lasting effect on the role and importance of the British pound or whether it will be a short-term shock to the currency. For example, in the case of the Swiss franc we observed that its pegging and unpegging has not permanently changed its reference status. Instead, it was a temporary change, and within one year of unpegging, the Swiss franc returned to cluster 3 where it was before the major interventions. However, the uncertainty induced by the pending Brexit negotiations with the European Union seems to leave this matter unresolved. While the pound spent only a total of 8 weeks in cluster 6, the cluster indicating the largest relative volatility, it since has remained in cluster 5 most of the time.

**JPY.** Despite its role as one of the IMF reference currencies, the Japanese yen is found at the tails of the market which correspond to relatively large movements. Unlike most other currencies, which tend to stay in the same cluster for extended periods of time and move slowly in Fig. 3, the Japanese yen appears to switch clusters rather frequently, moving mostly between cluster 5 and 6 starting in the year 2007.

It is noteworthy to highlight that this is not an artifact of the size of the interval we have selected to estimate the symbolic performance distributions, as this observation holds for longer and shorter time intervals. Instead the symbolic performance distribution changes frequently without deviating much from the cluster centers. We observe however that the central bank interventions trigger the move of the Japanese yen from cluster 6 to cluster 4, where it remains for a while. Likewise, it takes an outside shock for the yen to change its cluster association

[6]http://uk.reuters.com/article/uk-swiss-snb-idUKKBN15S1NU

[7]https://www.ecb.europa.eu/press/key/date/2012/html/sp120726.en.html

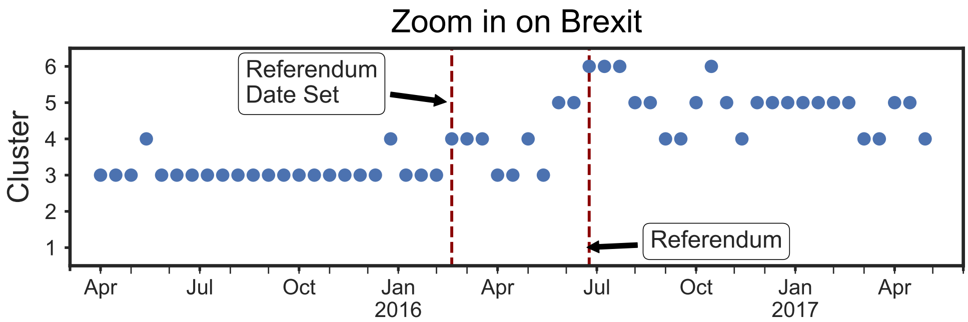


Figure 5: Cluster association of the British pound, highlighting the Brexit.

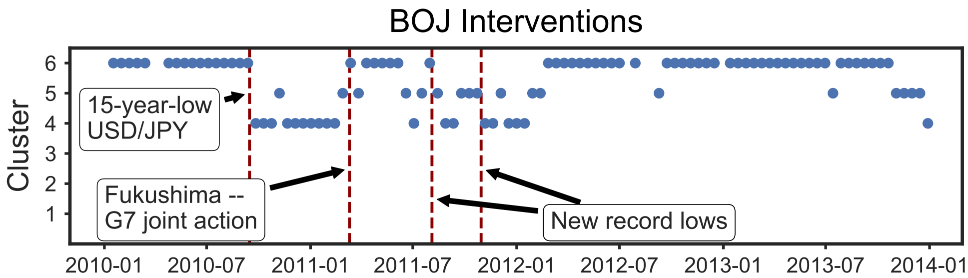


Figure 6: Cluster association of the Japanese yen, highlighting the time points and background of publicly announced central bank interventions.

again and return to cluster 6, as illustrated in Fig. 6. The first line marks an intervention on September 15, 2010, after reaching a 15-year-low of the dollar with respect to the yen, and it puts the yen in cluster 4 where it remains fairly steadily. The second intervention happened in the light of the Fukushima nuclear disaster. It is worth mentioning that this intervention was a joint effort by the G7, and thus it is not surprising that the yen returns to cluster 5.

The yen is mostly in cluster 3 in 2015; one may suspect that these changes are due to central bank actions. Given the USD/JPY rate development in 2015 and 2016, the Bank of Japan (BOJ) and government officials issued statements indicating the possibility of intervention. Our methodology indicates that these talks and any possible covert interventions have been successful, as the yen stabilized in its standard role in clusters 3 and 4 on the foreign exchange market in 2016. Recently, however, it has returned to clusters 5 and 6. Since Fig. 6 shows that the yen tends to eventually return to higher clusters a few weeks or months after interventions, we suspect that the BOJ has been intervening less actively or forcefully in the recent couple of years.

**MXN.** The Mexican peso has, until late 2011, mostly been classified into cluster 4 with some extended stints in clusters 3 and 5. During this time period the peso therefore can be characterized as a currency of average volatility when compared to the remaining currencies in our data set. Starting in late 2011, the peso shifts to cluster 5 and remains there for most of a time for a couple of years, reflecting a higher volatility. This coincides with growth rates which remained below expectations.

Since the Mexican peso is generally considered a satellite currency of the US dollar, political developments in the United States tend to be very influential. In fact, more than 90% of all peso transactions are trades on the USD/MXN pair. The 2016 presidential race has highlighted this strong connection, as Fig. 7 shows. The surge of Donald Trump in the primary polls of the Republican party in mid to late July 2016 corresponds to a change in cluster of the Mexican

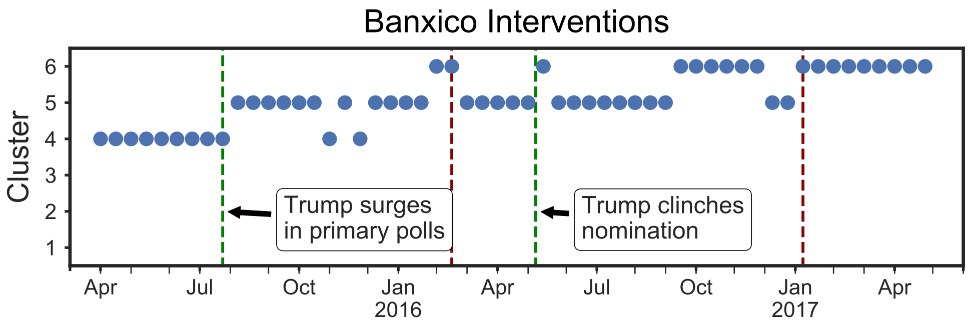


Figure 7: Cluster association of the Mexican peso, highlighting the time points of publicly announced central bank interventions in red and political events in the United States in green.

peso that moves from cluster 4 to 5. In early 2016, as Trump's nomination as presidential candidate became more likely, the peso moves to cluster 6. It only leaves this cluster to return to cluster 5 following a surprising intervention in February 2016 when Banxico intervened with a "shock and awe" move by selling US dollars and simultaneously increasing the interest rate[8]. The nomination of Trump as presidential candidate briefly sends the peso back to cluster 6. Since then, the peso has been in cluster 5 for some time and most recently in cluster 6.

**SGD.** Singapore's central bank is the Monetary Authority of Singapore (MAS). It is the mission of the MAS to hold the SGD in a managed float regime to control inflation in Singapore, which is detailed in biannual policy publications. Through direct intervention the MAS holds the Singapore dollar within a fixed band against an undisclosed trade-weighted basket of currencies. This method of controlling inflation stands in contrast to what most other central banks in our data set do, which is choosing to adjust interest rates to control inflation. Our clustering analyses show the results we would expect given the MAS policies and procedures: First, the Singapore dollar occupies cluster 3 for most of the time, interrupted by moves to cluster 2 and sometimes cluster 4. As the MAS enforces the policy band, bigger overall swings become less likely. This is typical of currencies in clusters 2 and 3. Furthermore, the main trading partners of Singapore comprise Hong Kong and China, whose currencies have been closely linked to the US dollar, and the United States. This implies that the cluster of the Singapore dollar tends to be similar to that of the US dollar. Figure 8 indicates with red vertical lines the times when the MAS publishes its policy updates, and we observe that changes in clusters for the Singapore dollar are often in accordance to these updates.

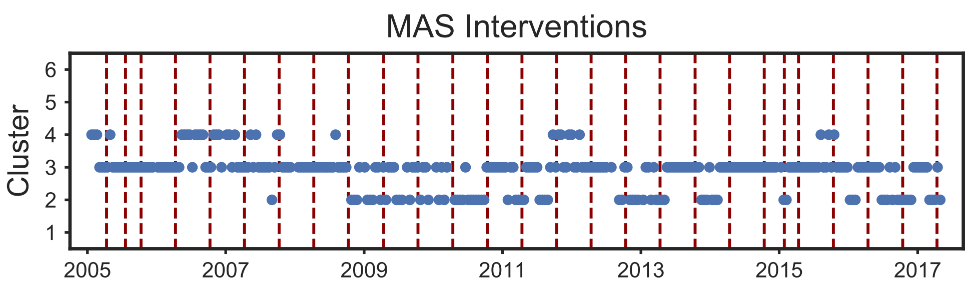


Figure 8: Cluster association of the Singapore dollar, highlighting frequent publications of policy briefs by the Monetary Authority of Singapore.

[8] https://www.ft.com/content/7129d13c-d364-11e6-9341-7393bb2e1b51

**USD.** For the majority of the time period the US dollar is associated with clusters 1 and 2 and appears to be a reference currency. In particular, the USD shares this role with the euro in cluster 1 during the year 2007, indicating very little movement with respect to the overall market. However, with the onset of the global financial crisis of 2008 the reference role of the US dollar within the foreign exchange market appears to weaken, as bigger swings and movements with respect to the remaining 13 currencies are more prevalent. This phase in which the dollar spends most time in cluster 3 lasts until late 2012. Hereafter the US dollar mostly stays in cluster 2 with some appearances in cluster 1 in 2014, reflecting the upswing of the US economy in the wake of the persisting European sovereign debt crisis and fear of deflation.

# 5 Conclusion

We have introduced a symbolic performance measure to quantify the roles of currencies in the foreign exchange market. Instead of investigating currency pairs, this novel approach allows for analysis of currencies individually, embedded within the entire system. We are able to identify the central (or reference) currencies in the foreign exchange market, and unsurprisingly the euro plays a major role, as do the US dollar and the British pound. Our methodology is also effective in observing the roles of currencies over time. We find that, in general, currencies maintain certain positions in the FX market for extended periods of time.

The independence from the choice of base currency in our methodology and its design enable us to closely investigate the characteristics of individual currencies with respect to the rest of the currencies in the network. We use our methodology to analyze the effects of central bank interventions, and we could hypothesize when a covert intervention by a central bank seems to have occurred. Furthermore we can examine the impact of specific shocks to the system, such as the pegging of the Swiss franc with respect to the euro, major interventions of the Bank of Japan, and the vote for Brexit.

We suggest that our methodology can be used to trace movements in the foreign exchange market to detect currency interventions, political developments and economic downturns or booms. One natural expansion of this work would be to include analysis of real-time data, monitoring the distribution of symbolic performances on shorter time scales and using observed turbulences in the foreign exchange market for currency forecasting.

### Acknowledgments

MK thanks Thomas Guhr for financial support for a research stay at the Center for Polymer Studies, Boston University. This work was partially supported by the National Science Foundation (NSF) [SES 1452061].

### Disclosure statement

The authors declare no potential conflict of interest.

# References

Aitchison, J. (1982). The statistical analysis of compositional data. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 139–177.

Aitchison, J., Barceló-Vidal, C., Martín-Fernández, J. A., and Pawlowsky-Glahn, V. (2000). Logratio analysis and compositional distance. *Mathematical Geology*, 32(3):271–275.

Almekinders, G. J. and Eijffinger, S. C. W. (1996). A friction model of daily bundesbank and federal reserve intervention. *Journal of Banking & Finance*, pages 1365–1380.

Baillie, R. T. and Osterberg, W. P. (1997a). Central bank intervention and risk in forward market. *Journal of International Economics*, 43:483–497.

Baillie, R. T. and Osterberg, W. P. (1997b). Why do central banks intervene? *Journal of International Money and Finance*, 16(6):909–919.

Beine, M., Bénassy-Quéré, A., and Lecourt, C. (2002). Central bank intervention and foreign exchange rates: new evidence from figarch estimations. *Journal of International Money and Finance*, 21:115–144.

Beine, M., Laurent, S., and Lecourt, C. (2003). Official central bank interventions and exchange rate volatility: Evidence from a regime-switching analysis. *European Economic Review*, 47:891–911.

Beine, M., Laurent, S., and Palm, F. C. (2009). Central bank forex interventions assessed using realized moments. *Journal of International Financial Markets, Institutions & Money*, 19:112–127.

Bonser-Neal, C. and Tanner, G. (1996). Central bank intervention and the volatility of foreign exchange rates: evidence from the options market. *Journal of International Money and Finance*, 15(6):853–878.

Cheng, A.-M., Das, K., and Shimatani, T. (2013). Central bank intervention and exchange rate volatility: Evidence from japan using realized volatility. *Journal of Asian Economics*, 28:87–98.

Dominguez, K. M. E. (1998). Central bank intervention and exchange rate volatility. *Journal of International Money and Finance*, 17:161–190.

Dominguez, K. M. E. (2006). When do central bank interventions influence intra-daily and longer-term exchange rate movements? *Journal of International Money and Finance*, 25:1051–1071.

Fatum, R. (2008). Daily effects of foreign exchange intervention: Evidence from official bank of canada data. *Journal of International Money and Finance*, 27:438–454.

Fatum, R. and Hutchison, M. M. (2002). Ecb foreign exchange intervention and the euro: Institutional framework, news, and intervention. *Open economies review*, 13:413–425.

Fatum, R. and Hutchison, M. M. (2003). Is sterilised foreign exchange intervention effective after all? an event study approach. *The Economic Journal*, 113:390–411.

Martín-Fernández, J. A., Barceló-Vidal, C., and Pawlowsky-Glahn, V. (1998). Measures of difference for compositional data and hierarchical clustering methods. In *Proceedings of IAMG*, volume 98, pages 526–531.

Tibshirani, R., Walther, G., and Hastie, T. (2001). Estimating the number of clusters in a data set via the gap statistic. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(2):411–423.