When Econometrics Met Information Theory: A Real-Time Approach to Track Yen Exchange Rate

Nan Wang

Abstract Fluctuations in currency exchange rates have enormous economic impact, affecting a wide range of market participants including governments, banks, corporates and individuals. Tracking currency exchange rates is crucial in many perspectives. While the financial press closely tracks and discusses major currency exchange pairs on a high frequency basis, scientific literature has been mostly silent in studying their patterns. This paper proposes an innovative approach in the endeavor of tracking the yen exchange rate against US Dollar (USDJPY), the second most traded currency pair. The proposed approach applies econometrics and information they in a real-time manner, leveraging the former on compressing high dimensional sparse data and the latter on quantifying nonlinear dependency. Using merely macroeconomics information and strictly avoiding look-ahead bias, the resulting tracking index from this approach achieves significant linear correlation with USDJPY and demonstrates ability to explain about 37% variances in USDJPY fluctuations over a 13-year period. The proposed approach has three phases: "vintage" generation, "vintage" storage and "vintage" transformation. The proposed approach is based on a deployed system which was started building in 2007. Since then, the system has been evolving systematically all the time to adapt to the ever-changing global macroeconomic environments.

The proposed approach is innovative in the following four aspects:
Use a real time database
Cover a broad range of macroeconomic information
Update information in a timely manner
Consider the dependency between countries.

Keywords Foreign exchange · USDJPY · Econometrics · Information theory · Dynamic system · Real-time database · Nonlinear dependency · Complex system

N. Wang (✉)
DeepMacro Inc, New York City, NY 10017, USA
e-mail: miya117517@gmail.com

D. Braha et al. (eds.), *Unifying Themes in Complex Systems X*, Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-030-67318-5_11

1 Introduction

Foreign exchange market is the largest and most liquid financial market in the world. According to 2019 Triennial Survey of turnover in OTC FX markets from the Bank for International Settlements, trading in foreign exchange markets valued at $6.6 trillion per day on average in April 2019. Foreign exchange market participants range from governments and banks (central and commercial), to commercial corporations, individuals and many others, all exposed to massive foreign exchange risk. Effectively tracking currency exchange rate is without doubt of high value.

Fluctuations of USDJPY particularly have tremendous economic impact compared to other currency pairs. According to aggregated Currency Composition of Foreign Exchange Reserves from IMF, the Japanese yen is the third most commonly used and the fourth most popular reserve currency as of 2019. Moreover, the yen has been a crucial indicator for economic activities. Through USDJPY, Investors express their views regarding the global economy, political event, fiscal policy and risk appetite [1]. For example, when macro climate is considered as risk-averse, the yen is usually aggressively purchased. During European sovereign crises in 2011, demands for the yen were unprecedentedly high, sending the yen to a historical high at 75.54 per Dollar [2].

Yet USDJPY patterns have received considerably less attention from scientific literature compared to its economic importance. One reason is that patterns of currency rates are notoriously difficult to track, and most academic models rarely work for even a bit longer than a brief period of time. Random walk is widely accepted as prediction benchmarks.

As a contribution, this paper proposes an innovative approach to effectively track USDJPY.

The approach is novel in four aspects.

Use a real time database. In the real world, macroeconomics data is changing all the time due to revisions.[1] The past looks different from today's "vintage"[2] than it did in the past. The proposed approach preserves all the data as they were at the time of release, and all the modeling are based on the latest published time series available at that point in time. Therefore, this approach effectively avoids information spillover, eliminates look-ahead bias,[3] or "forward looking bias" [3], and enables more reliable back testing.

Uses a broad range of economic information as building blocks. It's widely acknowledged that foreign exchange markets are primarily driven by macroeconomics factors. Yet macroeconomics covers many aspects. Besides, no single factor can explain a significant share of variances in exchange rate dynamics. Therefore, the proposed approach is aimed at covering as much economic information as possible. To name a few, information collected includes expenditure (i.e., retail

[1] See Sect. 2 for the definition of "revisions".

[2] See Sect. 2 for the definition of "vintage".

[3] Look-ahead bias occurs when using information or data in a study that would not have been known or available during the period being analyzed.

sales, household spending and passenger car sales), housing (i.e., condo inventories), Capex[4] (i.e., shipments of construction goods domestic use), exports, imports, labor (i.e., job offerings), industry (i.e., primary machinery orders), services (i.e., tertiary activity) and sentiment (such as consumer confidence). The source of the information could be either official, such as government agencies, central banks, and widely followed private sources, or unstructured alternative data[5] such as public internet displays, satellite images and social media. The benefits of integrating a broad range of information is to get earlier, orthogonal, and more detailed views of economic picture.

Building blocks are updated in a timely manner. Financial markets are pricing in various sorts of news, surprises or shocks all the time. Therefore, it is vital to keep the tracking dynamic and adaptive as well. In the proposed approach, the update frequency of input variables could be streaming, hourly, daily, monthly or quarterly, depending on the data availability. The proposed approach is designed to synchronize global economy, capturing and, if needed, factoring in new information at the first second.

The approach considers economic dependency between countries. Currency exchange rates are in relative terms only. Yet when modeling the driving force of exchange rate, common factors employed are individual-focused, such as economic growth [4], interest rate [5] and inflation [6]. Seeing the world as a highly intertwined network, we consider it important to take into account the dependency among countries and their interactions in terms of risk, trade and economic outlook when tracking their exchange rates. The proposed approach measures the business cycle dependency between Japan and the US in the efforts of tracking the relative positive of their currencies.

The rest of the paper is structured as follows. Section 1 contains definitions of terms referred in this paper. Section 2 covers the methodology of the proposed approach, discussing variables, econometrics, information theory metrics and the architecture of the proposed approach. Section 3 evaluates the experiment. Section 4 draws conclusions.

1.1 Definitions

This section provides definitions of the terms, including revisions, vintage, vintage point, reference point, real-time database, and alternative data, which are referred in this paper.

[4]Capital Expenditure.

[5]See Sect. 2 for the definition of “alternative data”.

1.2 Revisions

Revisions, in broad sense, are defined as a change in value for any reference point of the time series for a statistic when released to the public by an official statistical agency. Revisions can occur either when new observations (e.g. one additional month or quarter) become available and some past values are modified, or when the current and some past values are modified in an updated release of the current time series [7].

1.3 Vintage, Vintage Point and Reference Point

For a given time series we define vintage as the set of data (sequence of values) that represented the latest estimate for each vintage point in the time series at a particular moment of time. And that particular moment of time is considered as a reference point of the vintage [7].

1.4 Real-Time Database

A real-time database is a collection of historical vintages of the same time series. The revision on a given reference point for a time series can be identified in a real-time database as the change in value from an earlier vintage of estimates to a later vintage. A real-time database is popularly represented by a two-dimensional array, with vintage points in rows and reference points as columns. If the publication schedule of the time series has the same frequency as its periodicity (e.g. monthly), then the real-time database will have the appearance of a symmetric triangle as shown in Table 1 [7].

1.5 Alternative Data

Alternative data refers to data derived from non-traditional sources of information that might ultimately have investment alpha value.

2 Methodology

In this section, we describe the methodology for the proposed approach. We first describe input variables and their processing. Next, we introduce the employed econometrics model and the information theory metric. Finally, we describe the

Table 1 Example for a symmetric real-time database

Reference point	201,004	201,003	201,002	201,001	200,912	200,911	200,910	200,909	200,908
Vintage point									
201,004	−3.1298								
201,003	−4.7303	−3.7642							
201,002	−4.6134	−4.2548	−3.0574						
201,001	−4.8674	−4.5555	−4.3013	−3.2958					
200,912	−4.7807	−4.4307	−4.1613	−3.7999	−3.0997				
200,911	−4.7391	−4.3878	−4.0326	−3.6375	−3.0451	−2.5272			
200,910	−4.7998	−4.4742	−4.0248	−3.6214	−2.8978	−2.3338	−1.6303		
200,909	−4.8236	−4.5089	−4.078	−3.6165	−2.8787	−2.1536	−1.4258	−1.1593	
200,908	−4.8409	−4.5198	−4.0837	−3.6332	−2.8701	−2.0299	−1.1461	−0.7358	−0.7895

system architecture which contains three phases: vintage generation, vintage storage and vintage transformation. Phase one and two are implemented in a deployed factor system in DeepMacro, Inc (DFS). The system (DFS) was started building in 2007. Since then, the system has been evolving and improving by keeping bringing in alternative data, generating high frequency signals and applying nonparametric techniques.

2.1 Input Variables

Input variables are building blocks of the proposed approach. The DFS system is searching for candidates of input variables all the time from various sources. Each variable, whether sourced from an official or alternative source, goes through the following seven procedures before modeling.

Seasonal adjustment.

Conversion of nominal variables into real[6] ones.

Outlier and missing value detection.

"Ragged edge"[7] [8] solution (Kalman filtering [9]).

Missing early values solution.

Stationarity.[8]

Standardization.[9]

There are great challenges in applying these processes even to highly aggregated time series values from government sources. For some unstructured alternative data, challenges are magnified. For example, transforming satellite images into a normalized format requires paramount data reduction techniques.

Generally, there are five major categories of variables as described below.

Expenditure. This category covers four aspects: Capex, Imports, Housing and Government. Capex includes indices such as heavy trucks and shipment of computers. Imports include indices import aggregates from all major countries to the country in concern, or export aggregates from the country in concern to all major economies. Housing has indices such as new housing starts,[10] new home permits.[11] Government covers indices such as new school building construction and new healthcare building construction.

---

[6]The real value is the nominal value adjusted for inflation and other related measures, e.g. the nominal value of gross domestic product is usually adjusted by a deflator to derive its real values.

[7]Macroeconomics data is typically not available for today and the immediate past ("ragged edge") and subject to revision.

[8]A stationary time series is one with constant statistical properties such as mean, variance, autocorrelation, etc.

[9]Standardization is the process of scaling variables so that they are comparable.

[10]The number of privately-owned new houses on which construction has been started in a given period.

[11]Permits for new-home construction.

Income. This category covers one aspect: Labor. Labor includes indices such as new job offers, personal income, average hourly earnings, cash earnings scheduled, job offers/seekers ratio, unemployed, job opening vacancies and so forth.

Output. This category covers two aspects: Service and Industry. Service includes indices like real estate, wholesale, finance and insurance and transport. Industry includes indices such as manufacturers new orders, IP[12] motor vehicle and parts, IP headline, computers and capacity utilization rate total industry.[13]

Sentiment. This category covers three aspects: Household, Corporate and Financial. Household has indices such as conference board—present situation and consumer confidence (overall). Corporates have indices such as ISM non-manufacturing employment,[14] Richmond Fed[15] future 6 months, NAHB[16] next 6 months, Small business conditions index, Tankan[17] services for businesses and Economy Watchers Nonmanufacturing current.[18]

Others. This category covers all alternative data. Examples for this category are high-frequency industrial indicator processed from satellite images, hourly web traffic data and online job postings.

2.2 Econometrics: Dynamic Factor Modeling

The premise of a dynamic factor model (DFM) is that a few latent dynamic factors, f_t, drive the comovements of a high-dimensional vector of time series variables, x_t, which is also affected by a vector of mean zero idiosyncratic disturbances, e_t. The latent factors follow a time series process, which is commonly taken to be a vector autoregression (VAR). In equations, the dynamic factor model is,

$$x_t = \lambda(L) f_t + e_t \tag{1}$$

$$f_t = \psi(L) f_{t-1} + \eta t \tag{2}$$

where there are N series, so x_t and e_t are $N \times 1$, there are q dynamic factors, so f_t and η_t are $q \times 1$, L is the lag operator, and the lag polynomial matrices $\lambda(L)$ and

[12]Industrial production.

[13]The percentage of resources used by corporations and factories to produce goods in manufacturing, mining, and electric and gas utilities.

[14]The ISM Non-Manufacturing Index (NMI) is an economic index based on surveys of more than 400 non-manufacturing firms purchasing and supply executives, within 60 sectors across the nation, by the Institute of Supply Management (ISM).

[15]Federal Reserve Bank of Richmond.

[16]National Association of Home Builders.

[17]Short-Term Economic Survey of Enterprises in Japan.

[18]The Economy Watchers Current Index measures the current mood of businesses that directly service consumers, such as barbers, taxi drivers, and waiters.

$\psi(L)$ are $N \times q$ and $q \times q$. The ith lag polynomial $\lambda_i(L)$ is called the dynamic factor loading for the ith series, x_{it}, and $\lambda_i(L) f_t$ is called the common component of the ith series. All the processes in (1) and (2) are assumed to be stationary. The idiosyncratic disturbances are assumed to be uncorrelated with the factor innovations at all leads and lags.

DFM fits very well with macroeconomics, not only because of its ability to deal with sparse high dimensional data, but also its premise of latent factors driving co-movements of high dimensions. As forcefully argued by Lucas [10] and many others, the business cycle is not about any single economic signal but the dynamics and interactions of many signals.

The proposed approach uses DFM to extract a one-dimensional latent factor from all the input economic variables. The resulting latent factor is interpreted as factor of business cycles, or business cycle factor.

2.3 Information Theory: Transfer Entropy

Let $X = (x1, x2, \ldots, xN)$ and $Y = (y1, y2, \ldots, yN)$ denote two time series of simultaneously measured scalar quantities. The state spaces X and Y are reconstructed by the method of time delay embedding in which the state (or embedding) vectors are formed by delayed "past" scalar observations of the time series. The m-dimensional embedding vector is defined as:

$$x_t^{m,\tau} \triangleq \left(x_t, x_{t-\tau}, x_{t-2\tau}, \ldots, x_{t-(m-1)\tau}\right)^T \tag{3}$$

$$y_t^{m,\tau} \triangleq (y_t, y_{t-\tau}, y_{t-2\tau}, \ldots, y_{t-(m-1)\tau})^T \tag{4}$$

m is the embedding dimension, τ is the time lag, and $t = n, \ldots, N$ with embedding window $n = (m-1)\tau$. Note that the state vectors $x_t^{m,\tau}$, $y_t^{m,\tau}$ are points in the m-dimensional spaces X and Y. Essentially, transfer entropy (TE) measures the deviation from the generalized Markov property:

$$p(x_{t+1}|x_t^k) = p(x_{t+1}|x_t^k, y_t^l) \tag{5}$$

p denotes the transition probability. k, l denote time lag, If the deviation from the generalized Markov process is small, then we can assume that the state of space y has little (or no) relevance on the transition probability of the state vectors of space x. However, if the deviation is large, then the assumption of Markov process is not valid. The incorrectness of the assumption can be quantified by the transfer entropy which is formulated as Kullback–Leibler distance [11] between $p(x_{t+1}|x_t^k)$ and $p(x_{t+1}|x_t^k, y_t^l)$:

$$T_{Y \to X} = \sum_{x_{t+1}, x_t^k, y_t^l} p\left(x_{t+1}, x_t^k, y_t^l\right) \log \frac{p(x_{t+1}|x_t^k, y_t^l)}{p(x_{t+1}|x_t^k)} \tag{6}$$

TE represents the information about future observations of x_t obtained from the simultaneous observation of past values of both x_t and y_t, after discarding the information about the future of x_t obtained from the past of x_t alone.

The proposed approach applies TE in order to measure the dependency of the US business cycle factor on Japan. As one of the most influential informational methods to detect directed influence between two stochastic time series [12], TE is chosen because of its three attributes or functionalities.

- Nonparametric.
- Asymmetric and directional.
- Able to measure nonlinear coupling effects.

2.4 System Architecture

The proposed approach has the following three phases. Phase One and Two are based on a deployed factor system in DeepMacro (DFS).

Phase One: Vintage Generation. As shown in Fig. 1, Phase One generates vintages of factors on a daily basis. The process starts with integrating input variables from various data sources. Size of data sources could be kilobytes, megabytes or gigabytes. Variables are updated at different frequencies (ranging from streaming to annually) but are normalized on a fixed frequency basis (daily). Normalization has

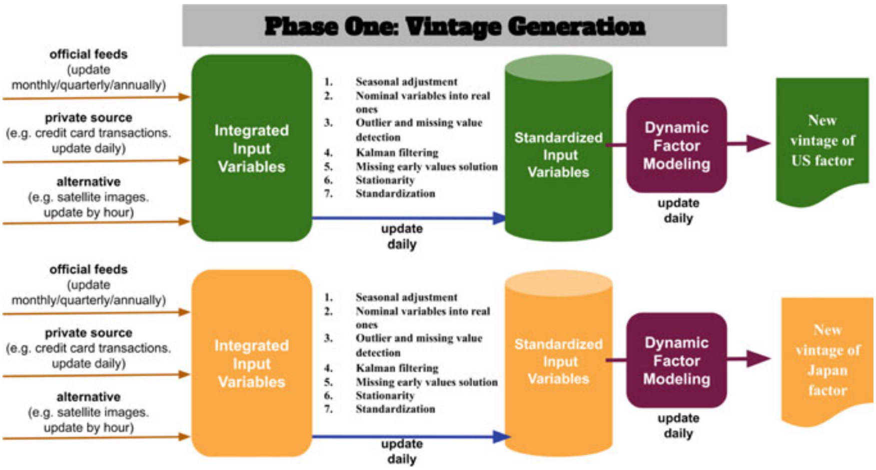

Fig. 1 Phase One, vintage generation

seven steps (as discussed detailly in Sect. 2.1), namely, seasonal adjustment, conversion of nominal variables into real variables, outlier and missing value detection, "ragged edge" solution, missing early values solution, stationarity and lastly standardization. After transformation, variables will be fed into dynamic factor modeling to generate new vintage of business cycle factors.

DFS updates business cycle factors every day, creating a vintage with every passing day as reference point. All vintages are monthly and 10 years' history long.

To give examples of business cycle vintages, Fig. 2 shows the 104 vintages of the US business cycle factor, with reference points being month-end dates between 28 Jul 2009 and 19 Mar 2018. In other words, with every month-end date between 28 Jul 2009 and 19 Mar 2018 being reference point, DFS computed a 10-year-long monthly time series up to the latest month available. As shown in Fig. 2, business cycle factor values on the same vintage points are different among vintages with different reference points. For example, the business cycle factor on Jan 31, 2005 vintage point varies about 1 standard deviation among vintages.

Phase Two: Vintage Storage. As shown in Fig. 3, every day whenever a vintage is created in Phase One, DFS stores it into a real time database where all the vintages are preserved.

Phase Three: Vintage Transformation. Phase Three transforms vintages of US and Japan factor and generates the USDJPY tracking index on a weekly basis (0:00 a.m. NY Eastern Time Every Monday). As shown in Fig. 4, every week, the latest vintages of the US and Japan factors are retrieved respectively from the real time database. Then, transfer entropy is applied on those two vintages with the dependency direction from Japan to the US, creating the USDJPY index.

Equations (6) and (7) illustrate Phase Three in detail. $V_{d_{(t)}}^{US}$ denotes a vintage of the US business cycle factor with reference point being $d_{(t)}$ and $V_{d_{(t)}}^{Japan}$ a vintage of Japan factor for reference point $d_{(t)}$.$d_{(t)m}$ denotes the latest vintage point for reference point d_t and $d_{(t)m-N}$ is the Nth latest vintage point. $F_{d_{(t)m}}^{US}$ represents the scalar value of US business cycle factor with vintage point being m and reference point being $d_{(t)}$.

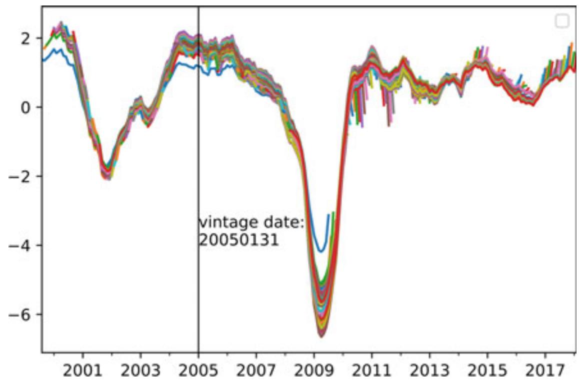

Fig. 2 Business cycle factor, US, 104 vintages with reference points being every month-end date, 28 Jul 2009–19 Mar 2018

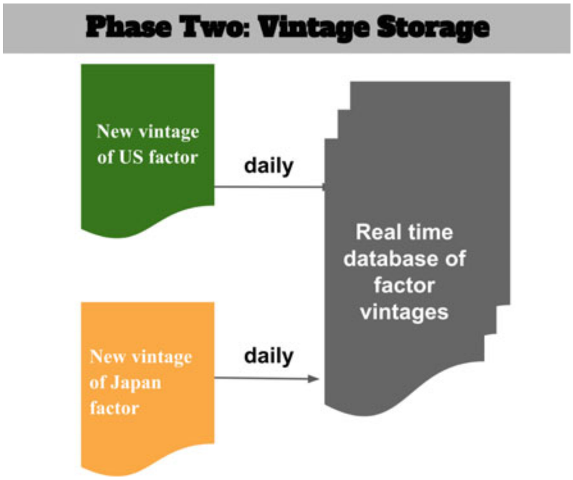

Fig. 3 Phase two, vintage storage

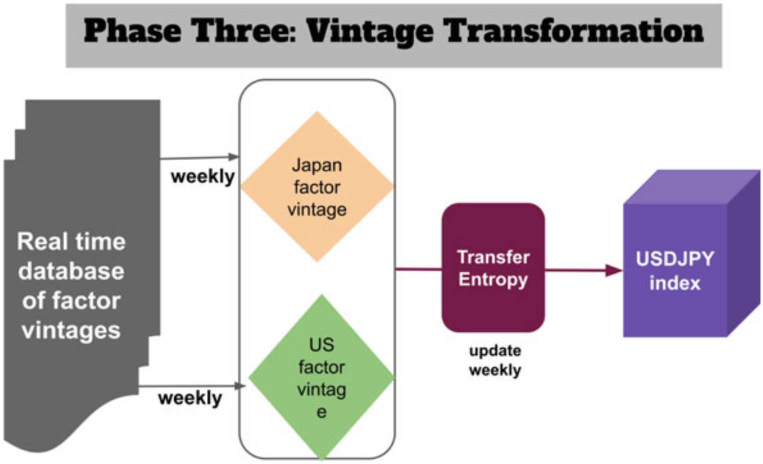

Fig. 4 Phase Three, vintage transformation

$$V_{d_{(t)}}^{US} = \left(F_{d_{(t)m}}^{US}, F_{d_{(t)m-1}}^{US}, F_{d_{(t)m-2}}^{US}, F_{d_{(t)m-3}}^{US} \cdots F_{d_{(t)m-N}}^{US}\right) \tag{7}$$

$$V_{d_{(t)}}^{Japan} = \left(F_{d_{(t)m}}^{Japan}, F_{d_{(t)m-1}}^{Japan}, F_{d_{(t)m-2}}^{Japan}, F_{d_{(t)m-3}}^{Japan} \cdots F_{d_{(t)m-N}}^{Japan}\right) \tag{8}$$

To calculate the USDJPY index for $d_{(t)}$, we apply TE on $V_{d_{(t)}}^{US}$ and $V_{d_{(t)}}^{Japan}$ with the dependency direction from Japan to the US. The resulting scalar, denoted as $T_{d_{(t)}}^{Japan \to US}$, is the US business cycle dependency on Japan on date $d_{(t)}$.

Similarly, we can generate a time series of $T^{Japan \rightarrow US}_{\{d_{t_1}, d_{t_2}, d_{t_3}, \ldots, d_{t_N}\}}$ for a group of specific vintage points $\{d_{t_1}, d_{t_2}, d_{t_3}, \ldots, d_{t_N}\}$. The resulting time series is the USDJPY index.

3 Evaluation

In this section, we describe our experimental setup of the proposed approach. We estimate its performance by evaluating its statistical correlation with the target.

3.1 Experimental Setup

Following the methodology described in Sect. 3, we generate vintages of the US and Japan business cycle factors and transform vintages into the USDJPY index.

For evaluation, we sample the US and Japan vintages of business cycle factors with reference points being all Mondays from 08 Jan 2007 to 30 Dec 2019.

3.2 Experimental Data

Business Cycle Vintages. 677 vintages of US and Japan business cycle factors respectively are used in our experiment to create the USDJPY index tracking USDJPY.

Figure 5 shows trending of the vintages of the US factor (upper) and Japan factor (lower).

USDJPY Index. The USDJPY index is generated from business cycle vintages. It is the proposed tracker for USDJPY. It is a one-dimensional time series with each scalar indicating the US business cycle dependency of Japan on a reference point.

USDJPY. USDJPY is a one-dimensional time series of daily closing (4.59 p.m. NY Eastern Time) spot US Dollar Yen exchange rate on selected reference points (Mondays, from 08 Jan 2007–30 Dec 2019)**.**

3.3 Experimental Results

We examine the results by looking at the correlation between USDJPY and USDJPY index. Figure 6 shows the correlation between the index and the target.

The index inversely correlates the dynamics of USDJPY in a significant linear way most of the time during the examining period. There exist some dates when the relationship deviating from the normal linearity.

Fig. 5 US (upper) and Japan (lower) Business cycle vintages. Reference points: Mondays, 08 Jan 2007–30 Dec 2019

From the scatter plot in Fig. 6 (upper), we can see that the deviation mainly happens during two periods, early 2011 to late 2012 and late 2014 to early 2015, both coinciding with the Bank of Japan unconventional monetary policies. During 14 Mar 2011 to 20 Dec 2012, the Bank of Japan announced enhancement of monetary easing nine times. On 31 Oct 2014, the Bank of Japan unexpectedly announced expansion of the Quantitative and Qualitative Monetary Easing, sending the Yen to its lowest value against the dollar in almost seven years.

An overall linearity in the examined relationship implies that the index is capable of closely tracking USDJPY in a significantly linear way. Brief deviations during abnormal monetary policy periods make sense considering that the index only factors in the dependency of macroeconomic growth.

To further validate the overall correlation, we exclude those two abnormal periods (14 Mar 2011–20 Dec 2012 and 31 Oct 2014–20 Apr 2015) and calculate the Pearson correlation between the index and USDJPY as well as Pearson correlation between their percent change transformations. Table 2 shows the result.

With the absolute value of coefficient for correlation being over 80% and that for the Pct. Chg. correlation being over 16%, together with significantly small p values, we infer that USDJPY index and USDJPY are highly correlated in significantly linear way.

Furthermore, we run an OLS (ordinary least squares) regression analysis. Table 3 shows the result. We can tell from the skew and kurtosis statistics that residuals are

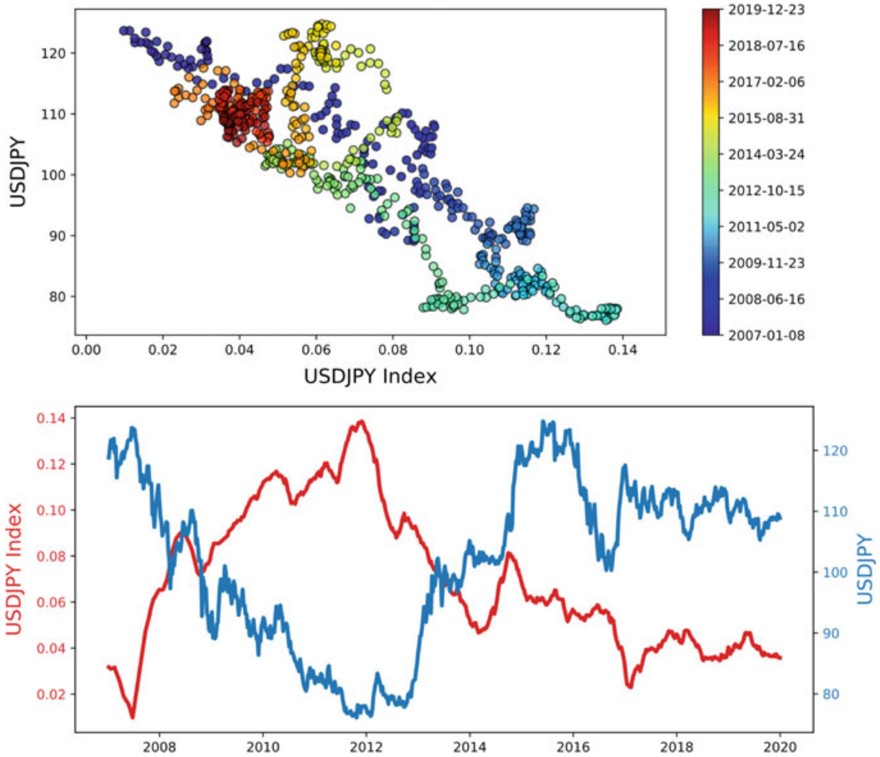

Fig. 6 Scatter Plot (upper) of USDJPY Index (x axis) and USDJPY (y axis). Colorized by dates. Jan 2007–Dec 2019. Trend plot (lower) of USDJPY Index (red) and USDJPY (blue), Jan 2007–Dec 2019

Table 2 Correlation between USDJPY index and USDJPY. Mondays, Jan 2007–Dec 2019. abnormal monetary policy periods excluded

	USDJPY USDJPY index	USDJPY Pct. Chg USDJPY Index Pct. Chg
Sample count	560	559
Correlation coefficient	−0.8081	−0.1693
P Value	<0.001	<0.001

fairly symmetrical, and a Durbin-Watson test statics of 1.9 indicates that the variance of the errors is independent of our independent variable here.

With reasonable amount of caution, we conclude that aside from abnormal monetary policy periods, the USDJPY index derived from mere macroeconomics information demonstrates ability of explaining about 37% variances in the US Dollar Japanese Yen exchange rates over a 13-year period.

Table 3 Regression analysis between USDJPY and USDJPY index

<table>
<tr><td colspan="3">Dependent variable</td><td colspan="4">USDJPY</td></tr>
<tr><td colspan="3">Independent variable</td><td colspan="4">USDJPY index</td></tr>
<tr><td colspan="3">R squared</td><td colspan="4">0.37</td></tr>
<tr><td colspan="3">F statistic</td><td colspan="4">209</td></tr>
<tr><td colspan="3">Prob (F-statistic)</td><td colspan="4"><0.001</td></tr>
<tr><td></td><td>Coef</td><td>Std Err</td><td>t</td><td>P > |t|</td><td>[0.025</td><td>0.975]</td></tr>
<tr><td>Const</td><td>119.124</td><td>0.946</td><td>125.879</td><td><0.001</td><td>117.263</td><td>120.985</td></tr>
<tr><td>X1</td><td>-228.720</td><td>15.805</td><td>−14.471</td><td><0.001</td><td>−259.803</td><td>−197.637</td></tr>
<tr><td colspan="7">Error analysis</td></tr>
<tr><td colspan="3">Skew</td><td colspan="4">0.430</td></tr>
<tr><td colspan="3">Kurtosis</td><td colspan="4">2.835</td></tr>
<tr><td colspan="3">Durbin-Watson</td><td colspan="4">1.930</td></tr>
</table>

4 Conclusion

We proposed a real-time approach for tracking the US Dollar Japanese Yen exchange rate. The approach is innovative in the sense that it uses a real time database, leverages a wide coverage of macroeconomics information, updates in a timely manner and considers the macro dependency between countries.

The approach has three phases with the first two based on a deployed dynamic factor system dated and the last phase creatively combining econometrics and information theory techniques.

Using macroeconomic information only and with strict forward-looking avoidance, our approach achieved a tracking index able to explain about 37% of the variances of USDJPY dynamics in monetary policy normal days over a 13-year-old period.

We expect the future work to achieve similarly significant results for other major currency pairs, such as EURUSD and GBPUSD, using the same methodology.

References

1. Botman, D., Danninger, S., Schiff, J.: Abenomics: lessons learned from two decades of conventional and unconventional monetary policies. In: Can Abenomics succeed? Overcoming the legacy of Japan's lost decades. IMF (2015)
2. The New York Times. https://archive.nytimes.com/query.nytimes.com/gst/fullpage-9C02E7DA153AF930A15753C1A9679D8B63.html. Last accessed 23 October 2011
3. Haselton, M., Nettle D., Andrews, P.: The evolution of cognitive bias. In: The Handbook of Evolutionary Psychology, pp. 724–746. Wiley, New Jersey (2005)
4. Ito, T., Isard P., Symansky S.: Economic growth and real exchange rate: an overview of the balassa-samuelson hypothesis in Asia. In: Ito, T., Krueger, O (eds.) Changes in Exchange Rates

in Rapidly Developing Countries: Theory, Practice, and Policy Issues, pp. 109–28. University of Chicago Press, Chicago (1999)
5. Sarac, T., Karagoz, K.: Impact of short-term interest rate on exchange rate: the case of Turkey. In: Murat Karagoz, M., Bildirici, M. (eds.) Procedia Economics and Finance 2016, vol. 38, pp. 195–202. https://doi.org/10.1016/S2212-5671(16)30190-3 (2016)
6. Kwofie, C., Ansah, R.: A study of the effect of inflation and exchange rate on stock market returns in Ghana. Int. J. Math. Math. Sci. https://doi.org/10.1155/2018/7016792 (2018)
7. McKenzie, R., Gamba, M.: Data and metadata requirements for building a real-time database to perform revisions analysis. In: OECD/Eurostat Task Force on "Performing Revisions Analysis for Sub-Annual Economic Statistics (2008)
8. Bouwman, K., Jacobs, J.: Forecasting with real-time macroeconomic data: the ragged-edge problem and revisions. J. Macroecon. (2011)
9. Harvey, A.: Forecasting, Structural Time Series Models and the Kalman Filter. Cambridge University Press, Cambridge (1989)
10. Lucas, R.: Understanding business cycles. Carnegie-Rochester Conf. Ser. Publ. Policy **5**, 7–29 (1977)
11. Kullback, S., Leibler, R.: On information and sufficiency. Ann. Math. Stat. **22**, 79–86 (1951)
12. Ito, S.: Backward transfer entropy: Informational measure for detecting hidden Markov models and its interpretations in thermodynamics, gambling and causality. Sci. Rep. **6**, 36831 (2016)