



**DUKASCOPY**  
RESEARCH PRODUCTS

26/09/2012



**Economic Research**

# Currency Pairs' Returns: Probability and Dependence

*The Dukascopy Bank SA research department starts a series of research on statistical properties of currency pair returns. In several issues we will investigate different characteristics to answer two main questions. Firstly, can we confirm or refute general assumptions about the nature of returns? And secondly, how do our findings affect the real-life use of the theory?*

In the current issue we focus on the relation between past and present returns and address subjects of distribution and dependency. Associated concepts, such as normality and independence, are crucial for models in risk analysis, prediction making, etc. That is why it is important to know whether the necessary assumptions hold, and therefore be able to estimate the accuracy of methods used in investing and trading.

We have chosen to study four pairs, trying to involve both popular and diverse currency couples: EUR/USD, EUR/GBP, EUR/JPY, and USD/CAD.

## Methodology

To see whether a set of historical data forms a pattern with known characteristics, we test it for having a certain type of distribution.

Distribution of a variable is a function, which maps the value of the variable to the probability of it taking the said value.

Therefore, knowing the distribution of returns, we would be able to estimate a probability of an exchange rate to take any value at any time.

In this paper we concentrate on two types of distributions: normal and stable, - as they are frequently mentioned regarding currency pair returns.

Normal distribution is a special case of stable distributions. It serves as a basis for many financial models, but assumes a number of strict characteristics that real-life data does not always possess. Stable distribution, on the other hand, is more general and allows greater deflections. Because of that, however, it does not adhere to the requirements of some of the models. Therefore, if the normality of returns is disproved, the usage of certain methods becomes questionable.

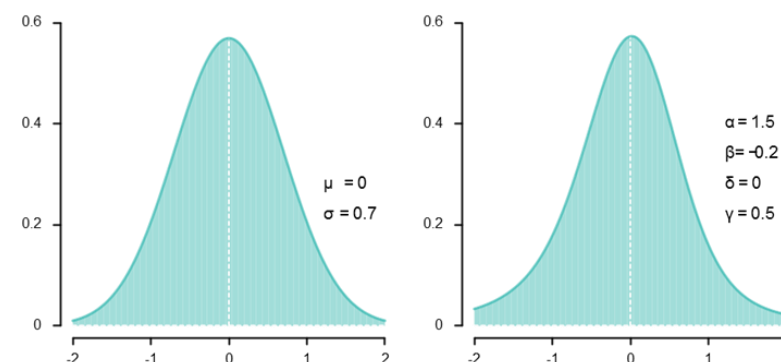


Figure 1 Examples of probability density functions for normal (left) and stable (right) distributions



We check the hypothesis for the distribution of returns by performing statistical tests. Two key factors for interpreting the results of such tests are significance level and p-value.

A statistical test is a procedure of making decisions about a population based on a sample. The statement under study (e.g. *returns are normally distributed*) is called a null hypothesis. A significance level is a probability of making a mistake and rejecting the null hypothesis when it is true. P-value is a numerical product received after performing the test. If the p-value is greater than the significance level, there are no grounds for rejecting the null hypothesis. Otherwise, the null hypothesis must be rejected.

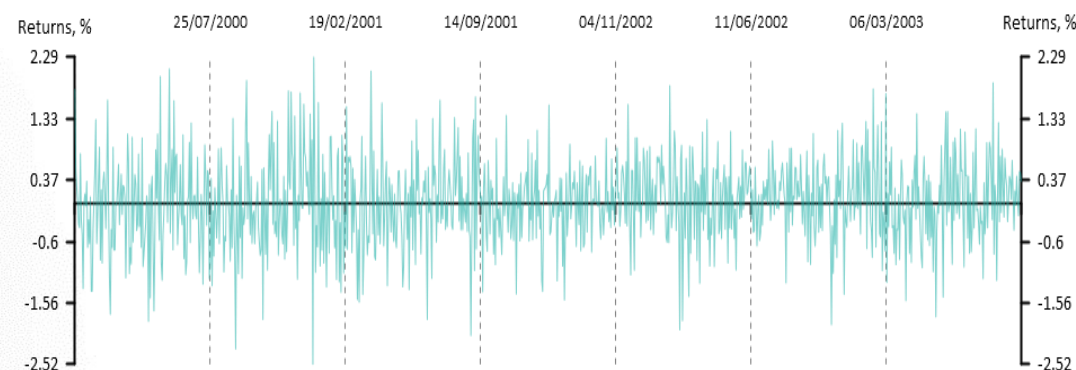
In this research we use five different statistical tests. The Kolmogorov-Smirnov (Lilliefors) (KS), the Anderson-Darling (AD), the Jarque-Bera (JB), and the Shapiro (Sh) tests are meant for testing hypotheses for distribution of returns. The Ljung-Box test is used to check whether returns in a sample are independent. More detailed information on the tests can be found in the Appendix.

Another way to investigate dependence among the returns is to run regressions.

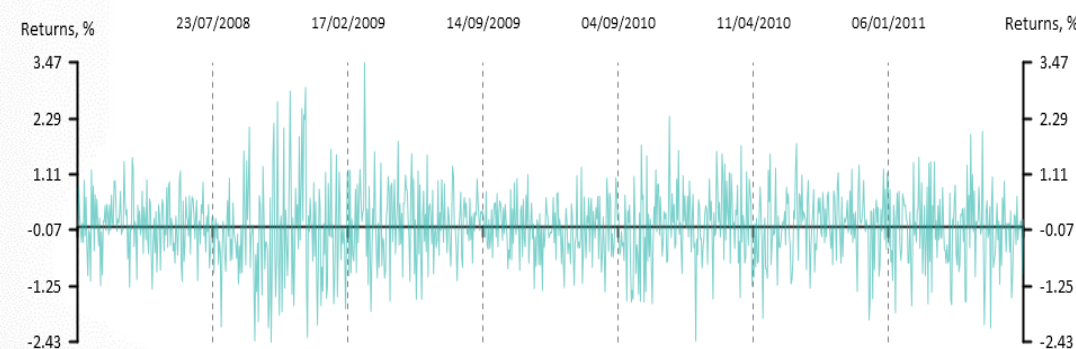
Regression is a method of estimating a relation between two variables, dependent and independent. It estimates the expected value of the dependent variable given a definite value of the independent one.

The dependent variable in our case is the return at a moment in time  $t$  (*current return*), while the return at  $t-1$  (*previous return*) is the independent one. This method gives us an opportunity to see whether returns are related to their predecessors, and thus inspect a possibility to predict future currency pair movements by the present ones.

The data used for testing are two sets of daily logarithmic returns, one collected over a time period from 01.01.2000 to 01.01.2004, and the other - from 01.01.2008 to 01.01.2012.



**Figure 2 EUR/USD logarithmic returns during 2000-2004**



**Figure 3 EUR/USD logarithmic returns during 2008-2012**

## Results

### 1. In general, logarithmic returns of the currency pairs are not normally distributed.

We have performed the tests for normality with a 0.05 significance level, and have received the p-values displayed in Tables 1-4. The red cells indicate cases in which the hypothesis about the normal distribution of returns does not hold, but the green ones – cases in which it cannot be rejected.

Firstly, we have considered a more widespread situation, when financial methods are applied to relatively small time periods. Thus we have divided our datasets into subsets that cover time periods of one year. It appears that the hypothesis of normality could be rejected for some samples and could not for others. In addition, there was no evident relation between p-values and time periods – the results changed chaotically, not showing any signs of gradual development. Two examples of different outcomes are presented in Tables 1-2. The results differ dramatically, from strong support of the hypothesis to its complete refutation.

	JB	KS	Sh	AD
EUR/USD	0.329	0.241	0.321	0.215
EUR/GBP	0.723	0.543	0.804	0.62
EUR/JPY	0.849	0.788	0.889	0.77
USD/CAD	0.637	0.515	0.462	0.366

**Table 1 P-values of Normality Tests (2000-2001)**

	JB	KS	Sh	AD
EUR/USD	0	0	0	0
EUR/GBP	0	0.007	0	0
EUR/JPY	0	0	0	0
USD/CAD	0	0	0	0

**Table 2 P-values of Normality Tests (2008-2009)**

	JB	KS	Sh	AD
EUR/USD	0.002	0.067	0.008	0.004
EUR/GBP	0	0.011	0	0
EUR/JPY	0	0	0	0
USD/CAD	0	0.007	0	0

**Table 3 P-values of Normality Tests (2000-2004)**

	JB	KS	Sh	AD
EUR/USD	0	0.014	0	0
EUR/GBP	0	0.004	0	0
EUR/JPY	0	0	0	0
USD/CAD	0	0	0	0

**Table 4 P-values of Normality Tests (2008-2012)**

We have then studied the full sets, assuming that a greater number of observations will allow us to get a more concrete result. The numbers in Tables 3-4 make it apparent that both sets consist of non-normally distributed returns, as all the tests produce p-values of 0 or very close to it. The only exception is the Kolmogorov-Smirnov test result for earlier EUR/USD returns. The test would also forbid rejecting the hypothesis for 2000-2004 EUR/GBP and 2008-2012 EUR/USD if the chosen significance level was 0.01, another popular threshold. As promising as these results are, compared with other p-values, they do not provide sufficiently strong evidence in support of normality. In fact, they say more about the specifics of the tests rather than the samples.



## 2. Currency pair returns are not always independent from their earlier selves.

Another important statistical property is data independence. In our case it holds if there are no bonds among returns in a sample. As was said before, the independence, too, can be checked with a statistical test, producing the already familiar p-values. The results for the 2000-2004 and 2008-2012 sets are given in Tables 5-6.

Lag	1	2	3	4	5	6	7	8
EUR/USD	0.077	0.14	0.198	0.131	0.098	0.157	0.22	0.146
EUR/GBP	0.115	0.277	0.063	0.025	0.032	0.048	0.01	0.01
EUR/JPY	0.388	0.455	0.55	0.68	0.576	0.557	0.636	0.718
USD/CAD	0.003	0.005	0.011	0.008	0.009	0.014	0.021	0.033

**Table 5 P-values of Ljung-Box Test (2000-2004)**

Lag	1	2	3	4	5	6	7	8
EUR/USD	0.192	0.308	0.343	0.384	0.521	0.551	0.588	0.641
EUR/GBP	0.055	0.02	0.032	0.065	0.101	0.097	0.114	0.169
EUR/JPY	0.667	0.418	0.474	0.457	0.426	0.326	0.256	0.315
USD/CAD	0.379	0.632	0.793	0.151	0.023	0.042	0.039	0.058

**Table 6 P-values of Ljung-Box Test (2008-2012)**

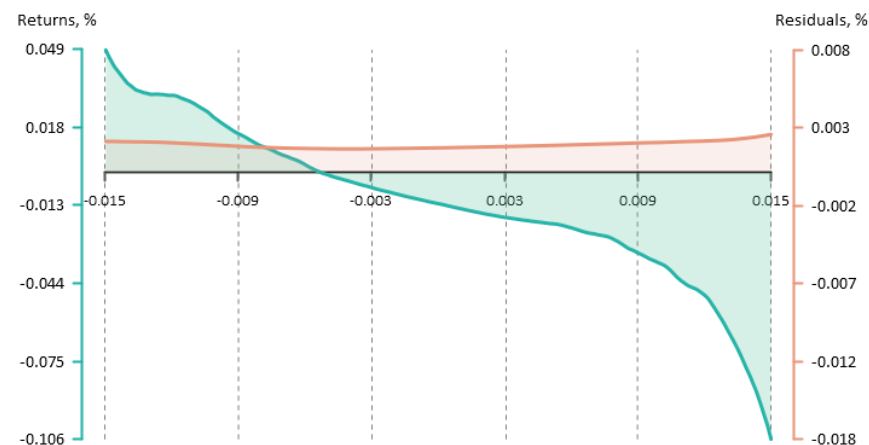
The results vary both among the currency pairs and throughout time. EUR/USD and EUR/JPY prove to have independent returns on any lag in both sets. For EUR/GBP the hypothesis gets rejected on certain lags, though not the same ones in different years. The 2000-2004 sample displays dependency on greater lags, starting from 4th, while the 2008-2012 – on 2nd-3rd, the lesser ones. This might indicate a change in the pairs' nature as recent movements are becoming more strongly related than the older ones. The most appreciable developments, however, are visible for USD/CAD results. The returns in 2000-2004 are dependent on all the lags, with especially strong evidence for the first ones, where the p-values are equal to zero. For the 2008-2012 sample, on the contrary, the hypothesis of independence holds for the first four lags, and gets rejected only on the last ones. Thus, in terms of independence, the pair has recently become appropriate for the application of financial and statistical models.

To take a closer look at the relationship between the lagged and non-lagged returns, we applied the regression analysis to the case with a one-day lag. The results are displayed as regression lines (green) and squared error lines (red).

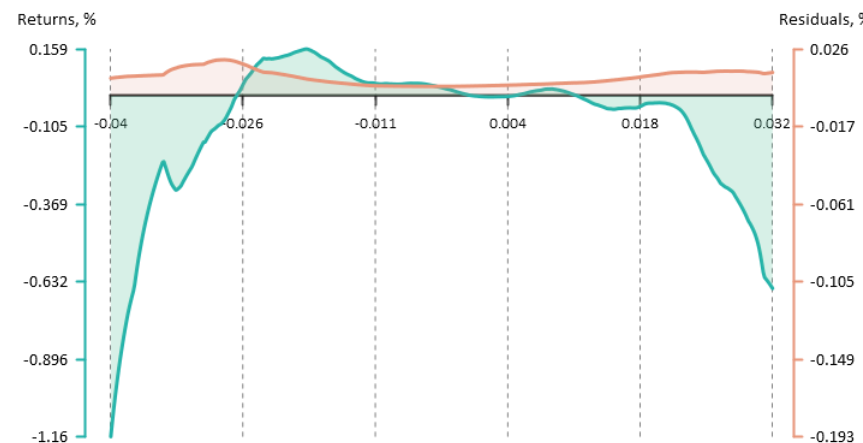
In general, as the test confirmed return independence from the first lag, the regression lines do not show any patterns in price changes. The only case in which the hypothesis of independence was rejected is 2000-2004 USD/CAD returns. This pair will thus be used to exemplify the difference in the nature of relationships. Results for other currency pairs can be found in the Appendix.

Figure 4 shows a clear, near-to-opposite relation between non-lagged and lagged returns. Here the best positive return is expected after the worst negative, and vice versa. Such a model describes zigzagging price movements. The errors of the model form an almost straight line on a low level, indicating a sort of stability. Though it might seem counter-intuitive, dependence in this case actually suggests an absence of short-term trends, with gains and losses constantly interchanging.

Figure 5, on the other hand, displays a regression for statistically independent data. We will now put aside the indicators of a worsening model (such as greater errors and unevenness of both error and regression lines) and concentrate on the suggested relationship. Surprisingly, even though the test for independence gave a clear positive result, the general form of the regression line does not differ much from the one on Figure 4. If we look at the horizontal values from -1.5% to 1.5%, we will see that the negative returns are expected to be followed by gains, while positive ones – by losses or neutral outcomes. The error in this interval is on approximately the same level, too. Thus the zigzagging nature of movements did not change, but possibly became more chaotic in its scale with volatility increasing over time. This might also be the reason for the change in the independence test results.



**Figure 4 USD/CAD logarithmic returns regression on lagged logarithmic returns with the residuals (2000-2004)**



**Figure 5 USD/CAD logarithmic returns regression on lagged logarithmic returns with the residuals (2008-2012)**



### 3. Distributions of returns are asymmetric and heavy-tailed.

As the tests gave strong evidence against the normality of returns, it is reasonable to try and estimate a different, better fitting distribution. We will focus on stable distribution, as its only requirement is independence of data, which in our case was confirmed by the test. The property of having a stable distribution is asymptotic, meaning it becomes clearer with bigger datasets. That is why we investigate the distribution parameters of the full sets of returns.

A stable distribution has four parameters, estimated values of which are given in Tables 7-8. Parameter  $\alpha$  describes the rate of decay of tails, i.e., the lower the parameter, the heavier the tails. For normal distribution  $\alpha=2$ . Thus Tables 7-8 show that the returns of all pairs take extreme values more frequently than they would if the normality held. In cases when  $\alpha>1$ ,  $\delta$  can be interpreted as a mean value. Here we see a confirmation of the belief that currency pair returns have a mean of 0. Parameter  $\beta$  characterises the skewness: its negative values indicate an extended left tail, but positive – a longer right one. Considering the 0 mean, estimated betas show that the returns tend to have more numerous and extreme negative values than positive ones. The only exception is USD/CAD, which seems to have started offering more gains in 2008-2012. Parameter  $\gamma$  shows how spread out the distribution is. However, it must not be confused with variance, which is infinite for all distributions with  $\alpha\neq 2$ . In our case it means that, theoretically, returns can take any values from infinitely negative to infinitely positive.

	EUR/USD	EUR/GBP	EUR/JPY	USD/CAD
$\alpha$	1.777	1.829	1.611	1.913
$\beta$	-0.052	-0.045	-0.111	-0.171
$\delta$	0	0	0	0
$\gamma$	0.0044	0.0035	0.0046	0.0027

**Table 7 Stable Distributions Coefficients (2000-2004)**

	EUR/USD	EUR/GBP	EUR/JPY	USD/CAD
$\alpha$	1.686	1.87	1.456	1.635
$\beta$	-0.267	-0.213	-0.174	0.236
$\delta$	0	0	0	0
$\gamma$	0.0049	0.0043	0.0051	0.005

**Table 8 Stable Distributions Coefficients (2008-2012)**

## Conclusion

We have investigated currency pair logarithmic returns, looking for patterns or laws that would describe price movements. The study showed that, in general, the returns are independent, heavy-tailed and skewed towards negative values. Thus the exchange rates move chaotically on a small time scale and often exhibit losses greater than gains. The returns also proved to have zero mean and infinite variance. All these properties result in the inaccuracy of normality-assuming models and consequently underestimation of risk, as strongly negative values appear more frequently than expected. Some samples, however, did demonstrate normality, proving that the properties of returns tend to vary for different time intervals and for different currency pairs. Thereby preliminary data testing is important, as some datasets are fitting for popular financial methods, but others require application of more complex and general models.



## Appendix

**Lilliefors test** – a modification of the Kolmogorov-Smirnov test for normality. Lilliefors tests the null hypothesis of data coming from a normally distributed population with unspecified mean and standard deviation. The Kolmogorov-Smirnov test, on the other hand, requires these parameters to be known. The alternative hypothesis for the Lilliefors test states that the population does not have normal distribution.

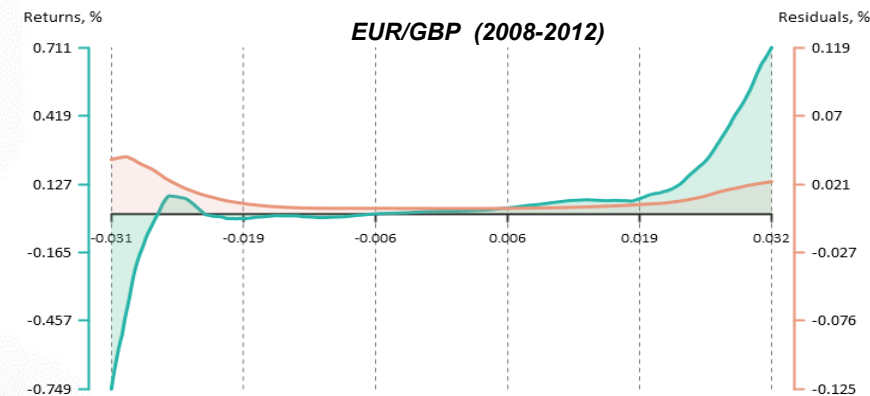
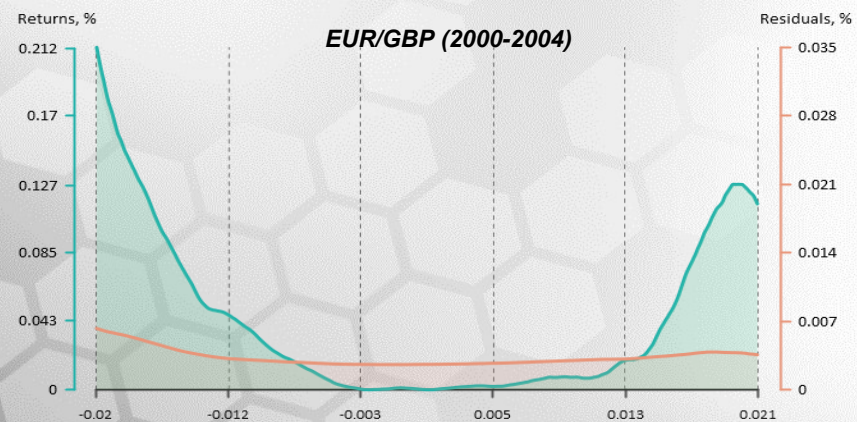
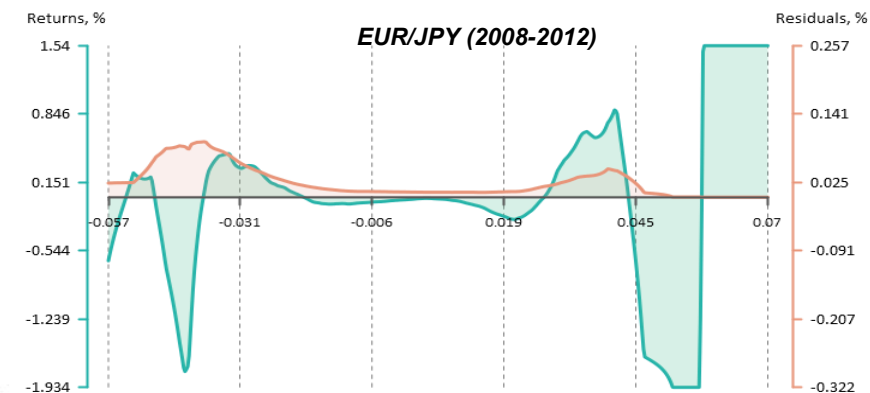
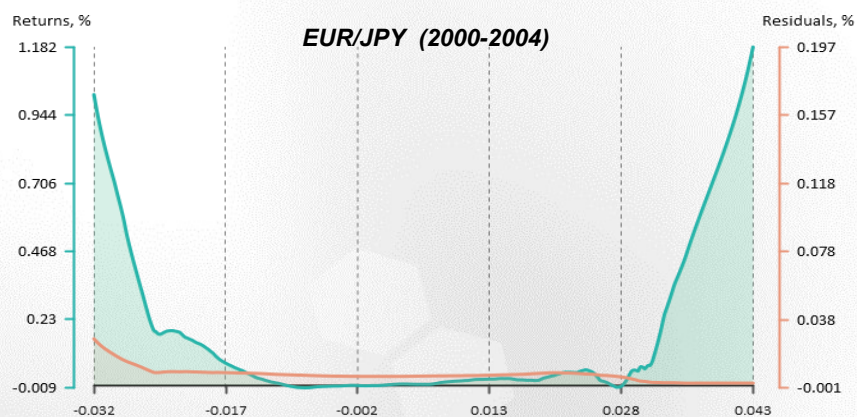
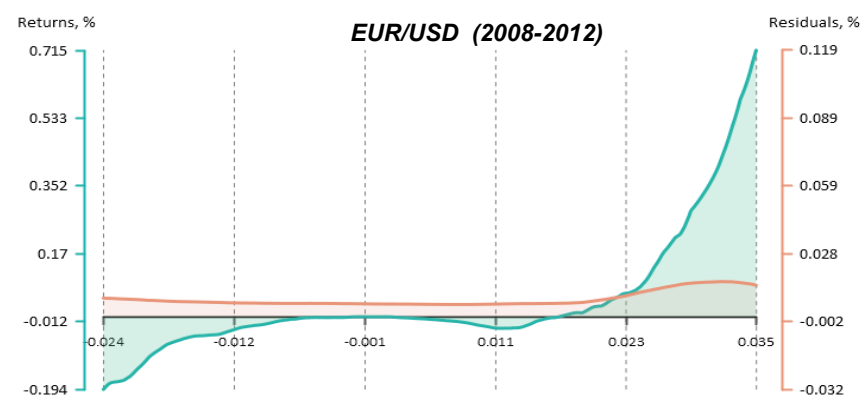
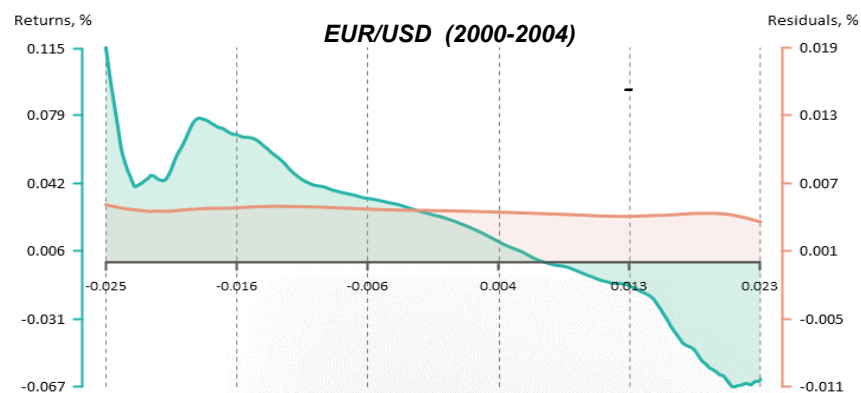
**Anderson-Darling test** – a modification of the Kolmogorov-Smirnov test that can be applied to different distributions. Therefore the null hypothesis suggests that data follows a specified distribution, but the alternative – that data does not follow that distribution. The Anderson-Darling test gives more weight to the tails of the sample than the Kolmogorov-Smirnov does, and thus is considered more accurate.

**Jarque-Bera test** – a statistical test that checks whether the sample's skewness and kurtosis correspond to the ones of normal distribution. Thus it tests the null hypothesis of the sample having zero skewness and zero excess kurtosis. The alternative hypothesis is that the data come from a non-normal distribution. The Jarque-Bera test is believed to be more accurate for large samples rather than for small ones.

**Shapiro-Wilk test** - a statistical test with the null hypothesis of population being normally distributed. It is often considered one of the most powerful normality tests, particularly for small samples. However, the accuracy results in the fact that rejecting the null hypothesis does not give any notion on which parameters of the population differ from normal.

**Ljung-Box test** – a statistical test on whether data are independently distributed. The null hypothesis is that the correlations within the population are zero, but the alternative – that the data are not independently distributed. This generalised alternative hypothesis gives an opportunity to check for a wide range of deviations from the main assumption. It does not specify the fault, but nevertheless shows whether the chosen model is in overall appropriate for the dataset.

## Logarithmic returns regression on lagged logarithmic returns with residuals







#### Newest releases and archive:

Fundamental Analysis  
Technical Analysis  
Press Review  
Market Research  
Expert Commentary  
Dukascopy Sentiment Index  
Trade Pattern Ideas  
Global Stock Market Review  
Commodity Overview  
Economic Research  
Quarterly Report  
Aggregate Technical Indicator

#### Additional information:

Dukascopy Group Home Page  
Market News & Research  
FXSpider  
Live Webinars  
Dukascopy TV  
Daily Pivot Point Levels  
Economic Calendar  
Daily Highs/Lows  
SWFX Sentiment Index  
Movers & Shakers FX  
Forex Calculators  
Currency Converter  
Currency Index  
CoT Charts

#### Social networks:



#### Disclaimer

Everything in this article, including opinions and figures, is provided for informational purposes only and may not be interpreted as financial advice or solicitation of products. Dukascopy group assume no responsibility for the completeness or the accuracy of any data contained in this article. Financial figures indicated in this article have not been verified by the Dukascopy group. Views, opinions and analyses are those of the author of the article, and are not endorsed by the Dukascopy group.

Dukascopy group waive any and all warranties, express or implied, regarding, but without limitation to, warranties of the merchantability or the fitness for a particular purpose, with respect to all information in this article. Dukascopy group shall under no circumstances be responsible for any direct, indirect, consequential, contingent or any other damages sustained in connection with the use of this article.