**Title: The influence of investor attention on the price of cryptocurrency.**

**Empirical Analysis of Investor Attention on Prices and Volatility of Bitcoin**

1 Influence of Investor Attention on Prices of Bitcoin

1.1 Variable descriptive statistics

1.2 Stability Test

1.3 Time Series Analysis

2 Investor Abnormal Attention Impact on Prices of Bitcoin

2.1 Variable Descriptive Statistics

2.2 Hausman test

2.3 Panel Data Regression

3 Investor Attention Impact on Volatility of Bitcoin

3.1 Descriptive Statistics

3.2 ARMA model establishment

3.3 GARCH Model Establishment

4 Results and Suggestions

5 Chapter Summary

**Empirical Analysis of Investor Attention on Prices and Volatility of Ethereum**

1 Description of Investor Attention on the Prices of Ethereum

2 Comparison of Investor Attention on the Prices of Ethereum

3 Influence of Investor Attention on Prices of Ethereum

3.1 Variable descriptive statistics

3.2 Stability Test

3.3 Time Series Analysis

4 Investor Abnormal Attention Impact on Prices of Ethereum

4.1 Variable Descriptive Statistics

4.2 Hausman test

4.3 Panel Data Regression

5 Investor Attention Impact on Volatility of Ethereum

5.1 Descriptive Statistics

5.2 ARMA model establishment

5.3 GARCH Model Establishment

6 Results and Suggestions

**Linkage Effect of Different Cryptocurrencies markets in Russia and China**

1 Study interval setting and sample selection

2 Research methods

2.1 Cointegration test

2.2 State space model

2.3 Quantile regression

3 Test of the Linkage of the Yield of Cryptocurrency markets in Russia and China

3.1 Unit Root and Cointegration Test

3.2 Granger causality test

3.3 Estimation of State-space Model of Cryptocurrency Returns Linkage under Emergent Crisis

3.4 Quantitative Regression Result of Cryptocurrency Returns Linkage in Emergent Crisis

4 Empirical Results and Discussion

## Data and proposed methodology

The methodological basis of the research was the methods of the system, complex analysis, abstract-logical, statistical, and functional, structural-level research methods, etc. The use of these methods allows us to substantiate the dynamics of the processes under study and integrate an approach that makes it possible to study the role of investment potential.

**Classification of methods of forecasting:**

Model prediction is a functional representation of adequately describing the process. Abbreviations of names as the models and methods. For example, there is a well-known model of auto-regression integrated moving average extended, ARIMAX. The model and the corresponding method are usually called ARIMAX, and sometimes the Box-Jenkins model on behalf of the sponsors.

The term "forecasting method" is a much broader concept of the "forecasting model". In this regard, in the first stage, the categories are intuitive and formalized.

**Intuitive forecasting methods** deal with judgments and estimates of experts. Today, they are often used in marketing, economics, politics, because the system, which is necessary to predict the behavior, which is very complex and not amenable to mathematical description.

**Formalized methods** are forecasting methods, which is to build a prediction model, define a mathematical relationship that allows you to calculate the future value of the process. It is necessary to proceed to the classification of forecasting models. In the first stage, the model should be divided into two groups.

**Domain models** are mathematical models of prediction, for the construction of which use domain laws. For example, the model on which make the weather contains equations of fluid dynamics and thermodynamics. Forecast of development of the population is based on differential equations. The forecast level of a person's blood sugar, diabetic, is based on a system of differential equations. In short, these models are used because of inherent in a particular subject area. This kind of model is peculiar to an individual approach in the design.

**Time series models** are the mathematical models that predict that you should find the dependence on the future values ​of the past in the process. These models are versatile for a variety of subject areas, that is, their general appearance does not change depending on the nature of the time series. We can use neural networks for the prediction of stock market indices.

**Time series analysis:**

Now we will consider more precisely this method. Time series analysis is a statistical technique that deals with time series data, or trend analysis. Time series data means that data is in a series of the particular time periods or intervals. The data is considered in three types:

**Time series data:** A set of observations on the values ​​that a variable takes at different times.

**Cross-sectional data:** Data of one or more variables, collected at the same point in time.

**Pooled data:** A combination of time series data and cross-sectional data.

Time series analysis is one of the main tools for observing collected data over time with sequential nature and predicting its future values.

### 

### Data

**Google Search Volume Index (SVI), Baidu Index, Deposit and withdrawal from active addresses of exchanges (Glassnode.com), Emission of new USDT coins (coinmarketcap.com), The complexity of the bitcoin network (Glassnode.com).**

Google Trends provides a search volume index (SVI) computed as the ratio of worldwide Google web search on specific keywords to the total number of Google searches over a given period. Moreover, we use Baidu Index to proxy the demand for cryptocurrency in each province in China. Baidu Index is a similar service with Google Trends, which provides Baidu query volume data from June 2006 to the present on a daily basis. Baidu search index reflects absolute numbers and is directly comparable. These data are normalized and scaled from 0 to 100 to make them comparable across regions. We download weekly data from January 2009 to modern days, will providing observations on each of cryptocurrency pairs: BTC/USD, ETH/USD, BTC/EUR, ETH/EUR. The choice of cryptocurrency pairs is based on their importance and the availability of SVI data.

The keywords we use in Google Insights are pairs of three-letter abbreviations for cryptocurrencies. For each cryptocurrency pair, we aggregate the Google SVI from the search in either order, for example, BTC/USD, ETH/USD, BTC/EUR, ETH/EUR. These abbreviations are unlikely to be subject to the problem of accidental increment in search volume, as in the case of SVI based on a firm’s ticker or name, both of which may have multiple meanings.

We will consider three measures of investor attention based on the Google SVI. The first is the level of attention index (hereafter “SVI\_level”), which corresponds to the original Google SVI. The second, the residuals from a regression of SVI level on monthly dummy variables (hereafter "SVI"), eliminates any seasonality. The third measure, the residuals from a regression of “SVI\_level” on monthly dummies and its own lagged values (hereafter SVI innovation), removes any seasonality and first-order autocorrelation. In the empirical analysis, we will report results for the seasonally adjusted attention measure, SVI.

In addition, widely used technical indicators reflecting the change in price over time (Stochastic Oscillator, MACD, Chande pulse generator, etc.). In addition, using the basic concepts of technical indicators in an attempt to identify more significant predictors, we form our primary data, obtained from databases, into additional functions. The second stage of preliminary data processing consists of two consecutive stages:

(a) data cleaning, which refers to missing and erroneous values.

(b) data conversion for the stringent requirements of some machine learning models, such as a neural network.