Continuous futures data series for back testing and technical analysis

Saulius Masteika[1+], Aleksandras V.Rutkauskas[2] and Janes Andrea Alexander[3]

[1, 2] Department of Finance Engineering, Faculty of Business Management, VGTU, Vilnius, Lithuania, EU;
[3] Center for Applied Software Engineering, Free University of Bozen, Bolzano, Italy, EU

Abstract. The main task of this paper is to present research results of specifying and adjusting historical data of futures market. We review and discuss the following methods for making continuous futures data series: Back/forward-adjusted, Proportionally adjusted, Gann, Perpetual, and Nth Nearest contract. The paper presents a comparison of historical data adjustment methods. The contribution of this research is to analyze graphically the methods and to find the least distortive techniques for different time frames and different trading strategies. We implemented this research and gave examples using historical data of the North American Futures market. The research results presented in this paper can be applied upfront for making back-testing of trading strategies in futures market based on technical analysis.

Keywords: futures; historical data; gap; continuous data series; technical analysis

1. Introduction

In 2010 the total number of contracts traded on futures exchanges around the world reached almost 11.2 billion, what makes a new all-time record and is 36.6% higher than in 2009 [1]. Because of increasing uncertainty in financial markets and growing demand for hedging businesses, trading volumes from equity and currency markets increasingly move to derivatives markets. The need for a deeper analysis of futures markets historical data is in a great demand.

Several basic approaches to analyse the markets exist. The first one is fundamental analysis, based on economic and statistical indicators. The second is technical analysis, based on the analysis of historical price changes [2]. Technical analysis is more related to short and medium term investments and is applicable for the determination of exact prices for opening or closing positions [3]. There is a difference between technical analysis in equity or currency and futures markets. The most complex and therefore requiring additional specifications and knowledge is the technical analysis of futures markets [4]. One difficulty with the futures historical data is associated with splicing contracts together at contract boundaries. There are several methods with many additional variables to create continuous futures data series [11][8][9][10]. These methods and possible influences on the results of back-testing strategies have to be analyzed in detail. The main contribution of this paper is the graphical and empirical research of different methods in the market for creating continuous futures data series.

2. Continuous futures data series

Futures contracts are standardized contracts between two parties to exchange a specified asset of standardized quantity and quality for a price agreed today but with delivery occurring at a specified future date [5]. Futures trade in series of short-lived contracts that are only active for a few months. For example the most active contract for *Russell 2000 mini (TF)* in July, 2011 is *TFU11*, i.e. Sept. 11. Around mid-September, most activity will roll over to the December contract *TFZ11* and afterwards the *TFU11* contract will expire.

The problem with historical data of the futures market arises when trying to calculate the long-term profitability of trading strategies based on technical analysis. The prices of different contracts of the same

---

[+] Corresponding author. *E-mail address*: saulius.masteika@vgtu.lt

underlying asset are different mainly because of contagno and backwardation factors [6][7]. Continuous futures data series must be generated in order to do long term back-testing qualitatively. For a better understanding the chart with unlinked data series are shown in the figure 1.
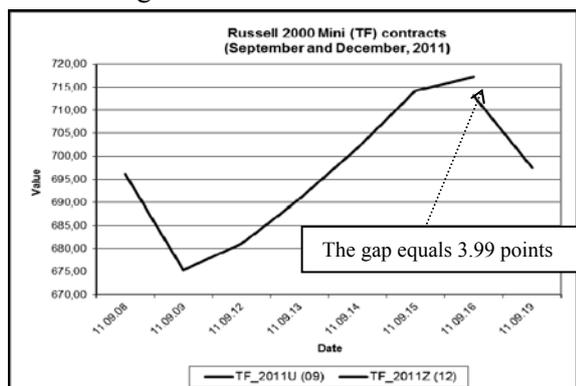

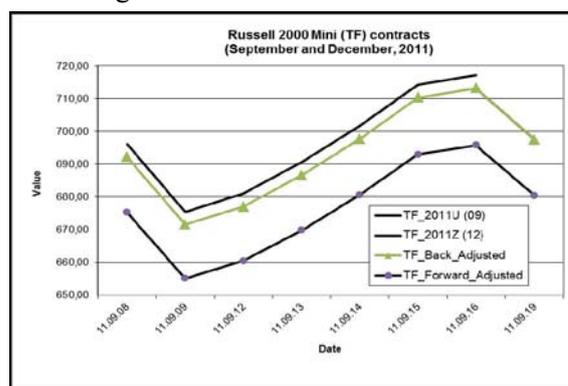Fig. 1 The gap in data series between following contracts   Fig. 2 TF back/forward- adjusted series chart

Fig. 1 shows that the gap (jump) trading *Russell 2000 Mini* contract is small, if compared to absolute values. But with the points converted into money it becomes obvious that the impact on profitability calculations is huge. The value of the gap equals to -$399 (-3.99 points x $100), just trading with one *Russell 2000 Mini* sized contract. In the calculation the multiplier 100 represents the full point value of the TF contract. The gap presented in Fig.1 would lead to inaccuracy and would distort the back-test and profitability calculations. The gap also would cause inaccuracies calculating numerical price indicators which are the basics of technical analysis. For example, a moving average will jump along with every gap and it will provide a false reading.

The solution to this problem is related with taking the gaps out of data series. Further in this paper a short overview of the most common methods to solve the gaps problem in futures data series is presented. After that method of minimum distortion on historical data is discussed.

2.1. Back/forward- adjusted series
Back-adjusted data uses the actual prices of the most recent contract with a backward correction of price discontinuities for successive earlier active delivery months. This adjustment requires to apply the raw price change of the earlier contract with respect to the price of the current (or later) contract [7][8][9]. For example, let's say we want the data series to be rolled backward through the quarterly contracts of June and September for Russell 2000 Mini futures (see Fig. 1). If the price of the June contract TF_2011U (06) is 717.19 and the price of September contract TF_2011Z (09) is 713.2 on roll-backward day (16/09/11), the method of back-adjusting would lower all prices for the June contract by -3.99. This would be maintained as a delta of -3.99 to be added to all the past data.

The same calculations go with forward-adjusted data series. Only the offset would be -3.99 points on the following, not the previous, contract. The advantage of this method is that it eliminates the need to recalculate old historical price data. The disadvantage is that forward-adjusted contracts do not represent actual values for today's markets. The data series with a back/forward adjustment for TF_2011Z (12) contract is provided in the figure 2. The chart lines in Fig. 2 show that the back-adjusted (TF_Back_Adjusted) data series is exactly the same as the current contract (TF_2011Z(12)) prices and at the same time the previous contract prices (TF_2011U(09)) are offset. Fig. 2 also shows that the application of forward price adjustment causes the futures contract prices (TF_Forward_Adjusted) to keep diverging further from actual prices as time goes on.

Generalizing back/forward- adjusted series, one could say that a back adjusting method could produce negative price values into the past as value and buying power of the money is decreasing. On the other hand, the forward-adjusted method could produce negative price values when a deflationary environment is present. This shortage is especially obvious when long term back-testing analysis is implemented. Most of the time the solution to this issue is achieved by adding a fixed amount of points making non-negative values or applying a proportionally adjusted data series.

2.2. Proportionally adjusted series

The proportionally adjusted method expects that the delta adjustment is done in percentage or ratio terms. The adjustment ratio is calculated by dividing the price of the new contract by the price of the old contract, which ensures a constant relationship between any prices across historical data. The main advantage of a proportionately adjusted series is that it cannot move into negative territory, instead it approaches zero. The weakness of the method is related with bigger data fluctuations, see Fig.3.
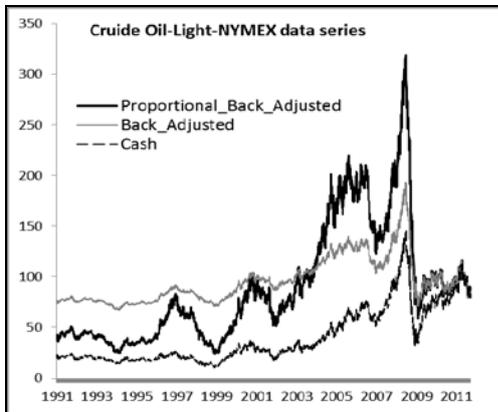

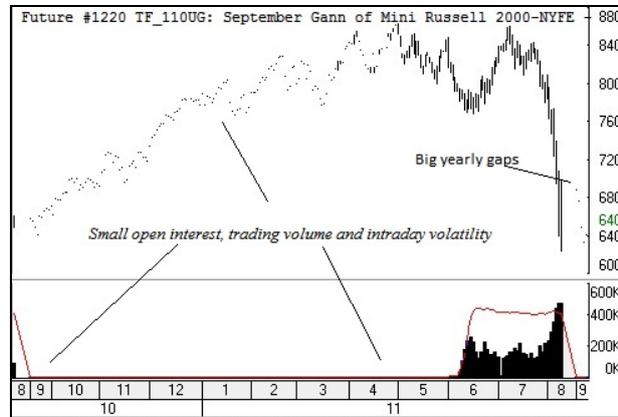Fig. 3 Proportionally_Back_Adjusted data series   Fig.4 Russell 2000 Mini Gann series chart (Chart by CSI UA)

The chart lines in Fig.3 show the differences between a proportionally back adjusted and a back adjusted data series. As we can see, the proportionally back adjusted data tends to move in a wider range. This is caused by applying an adjustment ratio when recalculating historical data. The adjustment ratio is calculated as percentage change. Therefore, a big difference between neighboring contracts causes bigger fluctuations in historical data series.

2.3. Gann series

The Gann series method transforms historical data into a series of successive historical segments of the same delivery month over successive years [10]. For example, using this method, the output of a continuous series becomes a compilation of all June deliveries. The example of Gann series chart with the TF_2011U(09) contract is presented in the figure 4. The chart in Fig. 4 shows that despite the advantage of offering a predictably long period of time to view a market on an annualized basis without adjusting the data, there is a big disadvantage of small open interest, trading volume, and intraday volatility. In the futures market, the total number of open long and short positions are equal. This total (long or short) is called the open interest. By definition, when a contract month first begins trading, the open interest is zero [11]. Just the last three months (June, July and August) are active in trading; for the Russell 2000 Mini September contract, see Fig.4. In summary, the Gann series method may be better than the Nth nearest-future variety because there are fewer discontinuities, i.e. jumps. The Nth Nearest contract is an artificial continuous data series representing a concatenation of successive contracts over time, reflecting the price, volume, and open interest of the Nth nearest future. The gaps (jumps) between contracts are left unchanged in this data series. The weakness of the Gann series is that one-year segments of a data series may be too long to yield meaningful information and it can be difficult to apply it in real trading, where the execution slippage on a trade can be too big because of a low volatility. Also a low intraday volatility and a rather static end of day price puts a limit for technical analysis based on chart patterns recognition like candle sticks or bar charts.

2.4. Perpetual series

The perpetual series method allows making a smooth blend from one contract price level to the next. Trying to make a continuous data series it is not always clear when one contract ceases to be a predominant contract on the board and when another takes over. This method attempts to alleviate that concern by displaying some type of average, usually a weighted average, and the weights of which change as contract importance gradually shifts [11]. As the contract expiration date approaches, a progressively smaller percentage of the current contract and a larger percentage of the new contract is used. The weights are put usually on open interest, volume or the time left till contract expiration. The following figure presents the

historical chart of DJIA contracts (2011) using the perpetual series method for building continuous data series:
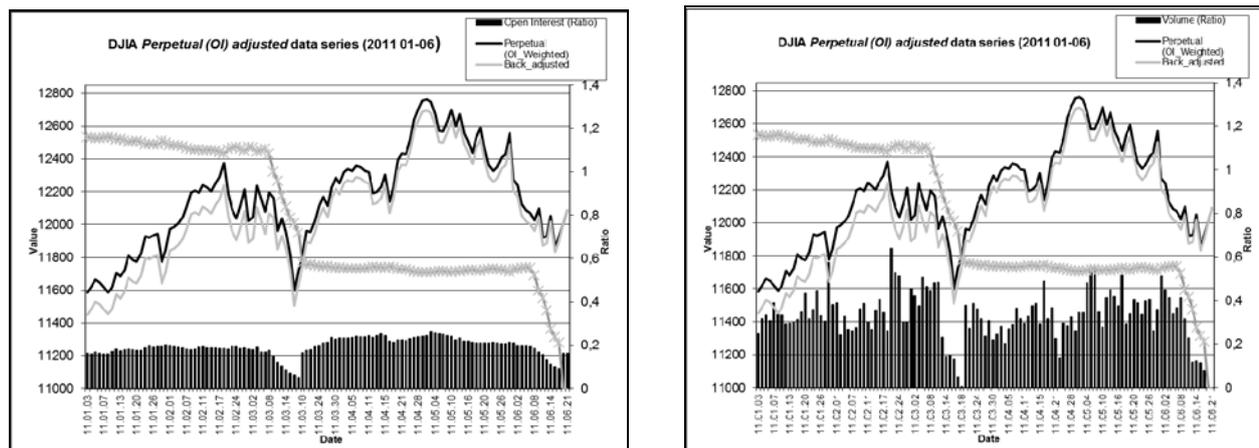

Fig. 5 Open interest- weighted perpetual data series

Fig. 5 shows the charts of DJIA (Dow Jones Mini) Back-adjusted and Perpetual (Open interest weighted) adjusted data series. In both charts the lines of Back_adjusted and Perpetual (OI_Weighted) smoothly converge closer as time goes on. The open interest is presented as a bar chart on the left chart of Fig. 5 and as volume on the right side. Both variables are represented on the second axis in the charts with ratio values. Fig. 5 shows that when the Open interest (Ratio) or the Volume (Ratio) is weakening, the Difference Ratio is lowering also. The difference ratio is a percentage difference between back-adjusted and perpetual (OI - Weighted) data series. This relation between Open interest and volume ratios and the Difference Ratio means that when the trading interest and volume is weakening, the contract's expiration date is approaching and a larger percentage of the new contract data is used. On the last day of the historical data taken (21/06/2011), the values using different adjusting methods come almost identical. Finally, we can say that the Perpetual method has an advantage when talking about statistical and smoothness issues but there is a big disadvantage because of non-tradable price representations. The method can be better applied as analytical tool but not as a tradable instrument.

3. Comparison of historical data adjustment methods

Which method is the best to build continuous futures data series and which one is the most applicable for back- testing and technical analysis is not an unambiguous question to answer. To make it easier, the advantages and disadvantages of different methods are presented in Table 1.

Table 1 Comparison of historical data adjustment methods

Method	Strengths	Weaknesses
Nth Nearest contract	No data adjustment. Represents actual values for today's market.	Gaps in historical data.
Forward adjusted series	Eliminates the need to recalculate old historical data. No gaps.	Does not represent actual values for today's markets. Possible negative price quantities into the past.
Back adjusted series	Represents actual values for today's markets.	Historical data must be recalculated. Possible negative price quantities into the past.
Proportionally adjusted series	Represents actual values for today's markets *(back adjusted)*. No negative price quantities when recalculating old historical data.	Historical data is recalculated. Designed for percentage based analysis methods. No absolute historical numbers.
Gann series	Long period of time in historical data on an annualized basis without adjusting the data.	The method can be difficult to apply in real trading, where the execution slippage on a trade can be too big because of low volume and open interest. Lack of OHLC historical data.
Perpetual series	Smoothens in data series. Good for statistical research.	Non-tradable price representations. Uncertain value of the smoothness ratio.

Which data series is the best for technical analysis depends on the time frame and parameters of the technical analysis indicators. For example, if back-testing is based on intraday trading and uses short term historical data – let's say up to three months – than Nth Nearest contract with no data adjustments is the best one to use. The same applies to Gann series when a rather short term (up to 4-5 months) technical analysis, e.g. short term chart patterns is performed.

On the other hand, if the technical analysis is based on long term indicators, e.g. moving averages (200 days), than eliminating gaps in the data series is necessary. Forward adjusted series can be used when back-testing trading strategies based on absolute level of past prices. Proportionally adjusted series can be best applied when the analysis is done on really long periods, like 50 years or longer. In such cases no negative price quantities are generated but an additional method of calculation is needed to have the ability of translating percentage based performance results into absolute values. The Perpetual method for making historical continuous data series is best applied when statistical analysis rather than technical is applied. In real trading, perpetual series are ineffective because of the non-tradable price representation. In real trading, placing exact orders in the market based on technical analysis back adjusted data series is the most applicable.

4. Conclusions

Recently, a new all-time high trading activity in futures exchanges but also lack of scientific considerations in the area motivated this research about historical data engineering of futures markets. The main goal of this paper is to illustrate the current research about futures historical data and to give some recommendations specifying, preparing and configuring historical data for the technical and back testing analysis. In summary, for statistical researches and cross market analysis the Perpetual series is the most suitable. For real trading activities where the real values of the market prices are necessary, Back and Proportionally adjusted series should be implemented. For short term technical analysis, Nth Nearest and Gann series could give most benefit. Generalizing we hope, that the results of the analysis presented in this paper could constitute as the resource for building computerized systems based on technical analysis.

5. Acknowledgment

6. References

[1]   W.Acworth,"Record volume 2010 (Annual volume survey)" Futures Industry, March 2011, pp. 12-29

[2]   M.N.Kahn, Technical analysis plain and simple: Charting the Markets in Your Language, 3rd edition, FT Press, 2010

[3]   S.Masteika, "Short term trading strategy based on chart pattern recognition and trend trading in Nasdaq Biotechnology stock market" Proc. Business Information Systems (BIS 10), Springer, May 2010, pp.51-57

[4]   R.Kolb, R.K.Overdahl, Understanding Futures Markets, Wiley-Blackwell, 2006

[5]   J.Liberty, "What everybody ought to know about continuous futures contracts", Automated Trading System, 2009

[6]   J. Hyerczyk, "The Data Game", Futures magazine, Aug. 2011.

[7]   B. Pelletier, Computed contracts: Their Meaning, Purpose and Application. An Essay on Computed Contracts // http://www.csidata.com/cgi-bin/getManualPage.pl?URL=essay.htm (referred on 17/09/2011)

[8]   B.Fulks, "Back-Adjusting Futures Contracts", Trading Recipes DB, 2000.

[9]   CSI, "Computed Contracts Get Real", Online Journal, July 2001// http://www.csidata.com/techjournal/csinews/200107/page01.htm#1

[10] Logical Information Machines (LIM), "Rollovers", Digital publication, 2008.

[11] J.Hyerczyk, "Continuous Data: Not so easy", Futures magazine, July, 2008.