Converting Forex Volume Into Dollars

A simple quantitative approach

The idea that volume is the missing link in Forex trading has been at the back of my mind for quite awhile. Despite persistent and sometimes vehement claims to the contrary, there is growing evidence that the tick volume we see on our Forex trading platforms is highly correlated with "real" volume across the entire FX spot market.

Back in about 2011 an ex bank trader by the name of Caspar Marney did some investigation and published his findings in a special supplement of FX Trader magazine. This article casts considerable doubt on the claim that Forex tick volume is useless. There is a link at the end of this article to Marney's original.

The dozen or so big banks that form the central hub of the Forex interbank market and account for the lions share of volume that gets traded, use the Electronic Broking Service (EBS) and/or the Reuters platform to access their network of market participants. The secrecy surrounding this data suggest to me that if the average retail trader was able to access it, a considerable trading advantage would be obtained. One of our many challenges as traders is to try and discover a way around this barrier of secrecy.

Being an ex bank trader, Marney was somehow able to get hold of EBS data as well as data from a number of other banking systems. Using this data, he was able to show that the daily variations in volume that we see on our MetaTrader and other retail platforms is highly correlated across the whole Interbank FX network.

This article will expand on some of Marney's ideas using a simple but often misunderstood statistical tool called correlation.

Correlation, simply put, is a measurement of the closeness of the relationship between two sets of numbers. Normally the sets are both the same size. Here is an example of two different sets

Data Set 1 : 1,2,3,4,5,6,7,8,9,10
Data Set 2 : 2,4,6,8,10,12,14,16,18,20

Both sets each have 10 numbers and we can see that the numbers in Data Set 2 are exactly twice the size of those in Data Set 1. These sets of numbers are not equal but because of their exact relationship to each other they are highly correlated.

Calculating correlations by hand is a rather tedious process. There are a number of software tools that can make the job easier. Probably the most commonly used is Excel. We can simply type in the data as shown in Fig 1 and enter the formula "**= CORREL(B3:B12,C3:C12**). Excel will then calculate the answer which in this case is 1 or 100%. These two datasets are 100% correlated. This means we can predict a value for Data Set 2 if we know the value in Data Set 1. This can be very useful in trading. If we can find something that is not only significantly correlated with the price we are interested in, but also has a leading relationship with that price, then we have the makings of a statistical advantage or edge.

	A	B	C	D	E	F
1						
2		Data Set 1	Data Set 2		Correlation	
3		1	2		= CORREL(B3:B12,C3:C12)	
4		2	4			
5		3	6			
6		4	8			
7		5	10			
8		6	12			
9		7	14			
10		8	16			
11		9	18			
12		10	20			
13						
14						
15						
16						
17						
18						

Fig1 A correlation calculation between two data sets using Excel

When working with market data, we also usually need to clean it. Missing or corrupt data elements can cause calculations to fail or produce incorrect results. For small data sets, we could do this with Excel but as the data gets larger and more complex, the task becomes increasingly difficult. There are a number of other statistical software programs available that are much more powerful than Excel and better suited to this sort of work. The one I mostly use these days is called “R”. It is an open source program that is absolutely free. It is well supported by a huge user community. There is also a free Integrated Development Environment (IDE) available for R called RStudio. The two work well together.

In mathematics, a picture is often worth a thousand words. I have used R to generate a graph of the correlation between the two data sets shown in Fig 1. Note the straight upwards sloping line with each data point right on the line.

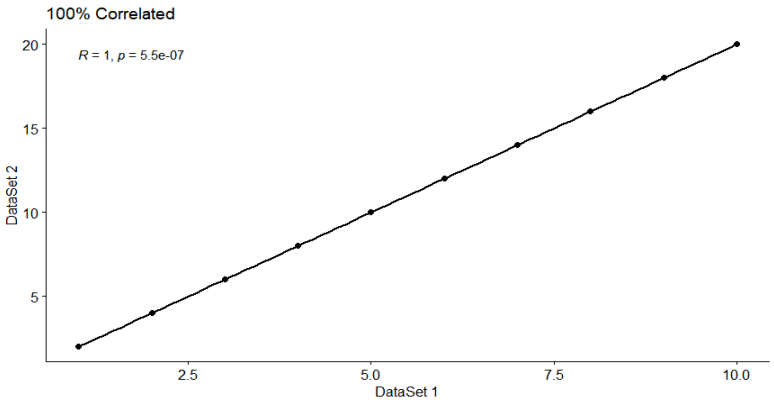


Fig 2.Graph of Data Set 1 and Data Set 2 showing a correlation of 1 (100%)

Data Set 1 : 1,2,3,4,5,6,7,8,9,10
Data Set 3: 2,7,12,2,14,5,18,12,2,10

If we take another Data Set 3 and test its correlation with Data Set 1 we get an altogether different picture. Now most of the points are well away from the line. These two data sets have a correlation of only 0.25 or 25% which is generally considered weak. This is the level of correlation we typically find in a trading environment with things like indicators if and when they are even working! When they are not working the correlations can fall to near zero. I have even seen traders using indicators that have a negative correlation where the price moves the opposite way to the predicted direction. If you have a reverse correlation and know about it you can often just reverse the process. If you don't know about it, you lose money. Generally speaking, any correlation much below about 0.3 or 30% is too weak to be of much use.

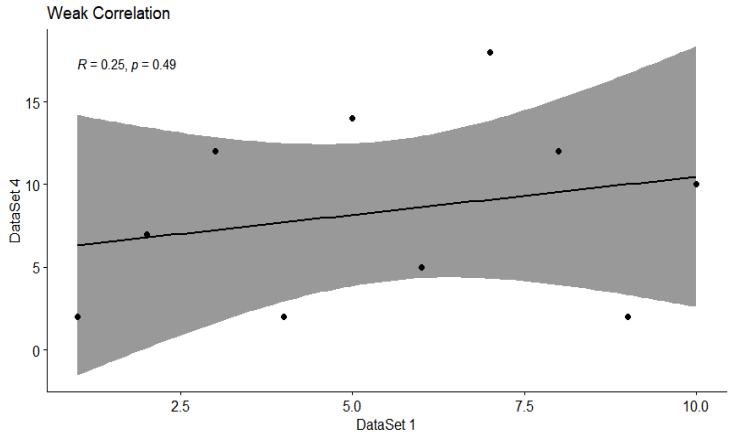


Fig 3. Graph of Data Set 1 and Data Set 3 showing a weak correlation of 0.25 (25%)

If we want to examine correlations between more than two sets of numbers there are easier ways to do it than trying to compare each set two at a time. Fig 4 shows the correlations between the five data sets on the left. These correlations range from very weak (6%) to very strong (100%). There are also some inverse correlations (negative values) which are opposites or reverse correlations.

Data Set 1 : 1,2,3,4,5,6,7,8,9,10
Data Set 2 : 2,4,6,8,10,12,14,16,18,20
Data Set 3: -2,-7,-12,-2,-14,-5,-18,-12,-2,-10
Data Set 4: 2,7,12,2,14,5,18,12,2,10
Data Set 5: 10,7,12,2,14,5,18,2,6,14

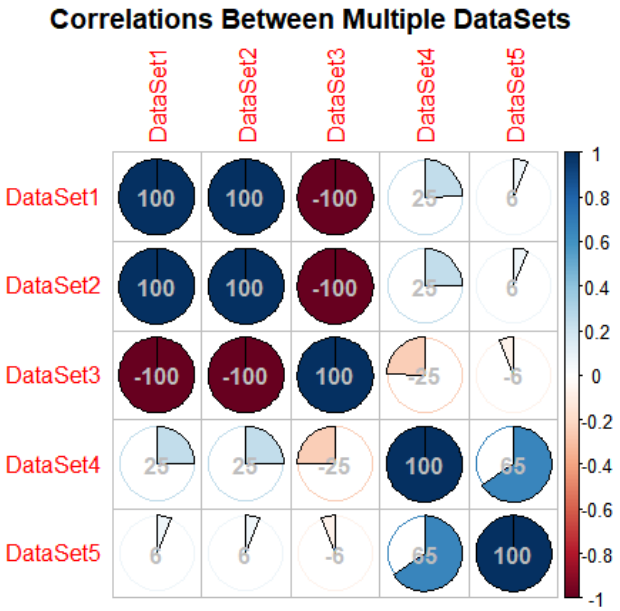


Fig 4 A correlation matrix showing correlations between 5 data sets.

Causal and Casual Correlations

A correlation between two sets of data does not by itself imply or prove that one thing **causes** another. This is a very common mistake that many people who use correlations make. Regular offenders of this important principle are the media and people with scientific and mathematical training who should know better.

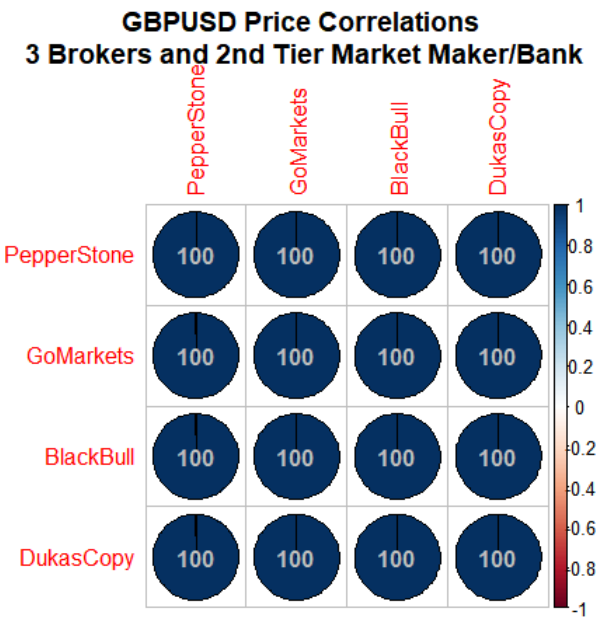
A typical example might be the discovery that people who live under high voltage power lines show a preponderance to develop certain types of cancer. This correlation does not **prove** that the power lines **cause** cancer. The correlation is **casual** until the hard work of developing, testing and proving a theory has been done. We also see this sort of error being made all the time in financial markets. A trading strategy based on correlations with proven market principles and behaviour is much more likely to be robust than one based on arbitrary casual relationships.

Spot Forex Price and Volume Correlations

I have chosen four different FX data providers to test for price and volume correlations. GoMarkets and Pepperstone are both well established retail broker networks with an international presence. BlackBull Markets is a small NewZealand based broker who appear to use a quite obscure liquidity provider. DukasCopy are a Swiss based 2nd tier market maker/bank. Hopefully the four have provided sufficient diversification for these studies.

The price correlation between all four was consistently high (> 99%) which was expected. If prices across the market varied, even for a very short period of time they would be be subject to arbitrage from alert and astute operators.

The correlations between tick volumes from the different operators was surprisingly high. These tick volume correlations ranged from 78% to 89% .



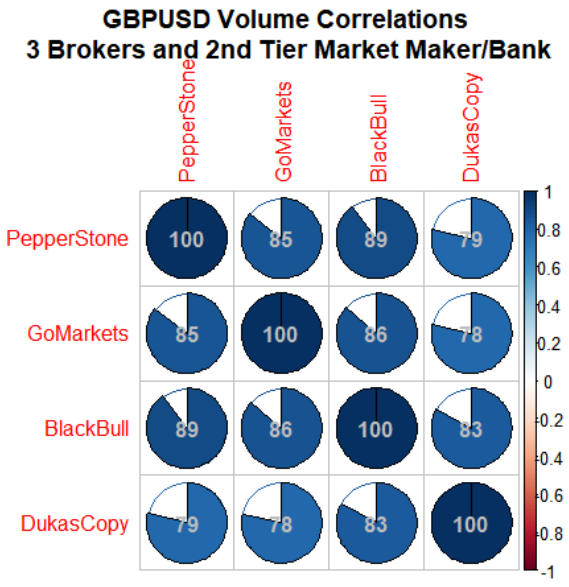


Fig 5 & 6 Correlation matrices showing for price and tick volume.

FX Futures data as a proxy for Spot FX

A number of traders use FX futures data with Depth Of Market (DOM) from the centralised Chicago Mercantile Exchange (CME) to help them make trading decisions. One of the questions that this enquiry attempts to answer is whether or not FX futures data provides a valid proxy for the FX spot market. They are after all entirely different markets and there appears to be no prima facie reason why either price or volume should be correlated. Fig 7 however shows that the prices between the two markets are highly correlated.

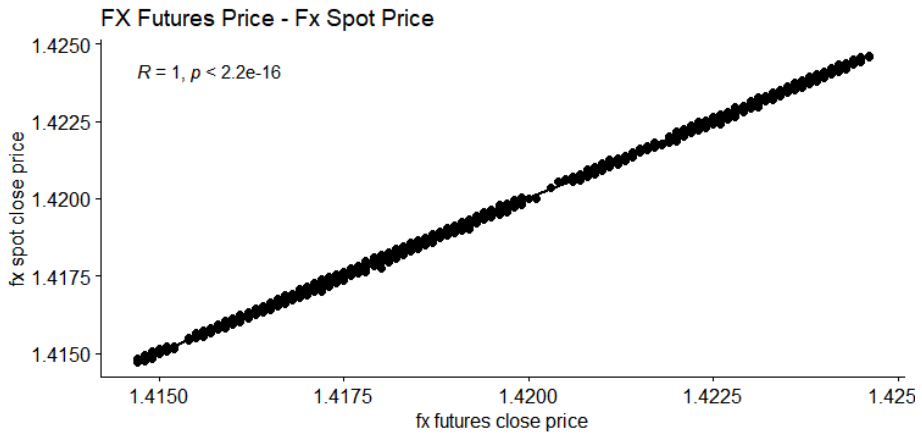


Fig 7. Fx futures price and Fx spot price show a very high correlation (>98%).

Correlations are also commonly used to determine if one data set has a leading or lagging relationship with the other. This can be very important in trading. It is typically advantageous to identify an indicator that leads price movement. The process involves moving one set of numbers relative to the other and testing the correlation at each step. This process is often referred to as cross correlation. Fig 8 shows a cross correlation for FX futures price with FX spot price. The highest correlation is achieved at 0 with no lag either +ve or -ve. According to this test neither price predicts the other. There have been reports that FX futures prices lead FX spot prices by a few seconds. This particular test only has a resolution of 1 minute so it would not detect such a small lead or lag. A more sensitive test would be required but such a test would also be very prone to being influenced by data noise.

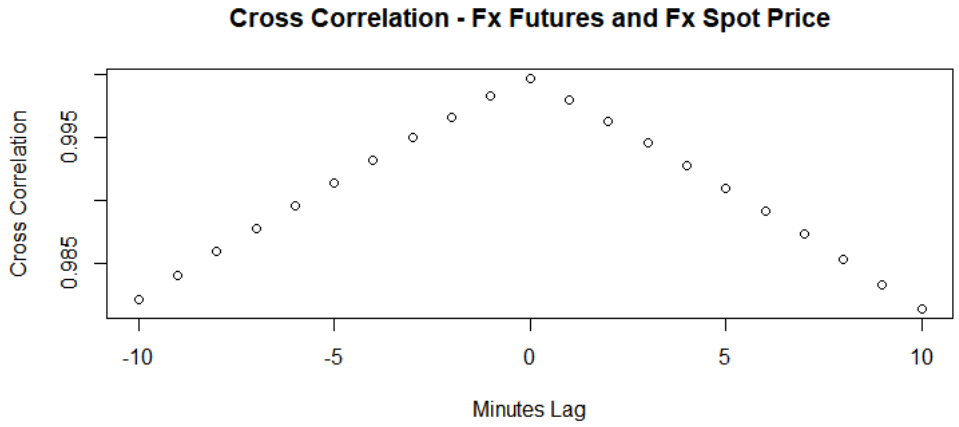


Fig8. A cross correlation test shows that neither spot or futures price leads or “predicts” the other .

The correlation graph between FX tick volume and FX futures volume in Fig 9 shows quite a different picture to the one for price correlation. We have already established that FX tick volume seems to be highly correlated with "real" volume in the spot FX market so the correlation between FX spot tick volume and FX Futures volume is a very important one. The correlation of 0.57 or 57% lies in the medium range. This suggest that while using FX futures data as a proxy for the FX spot market has some validity, the relationship is not a particularly strong one.

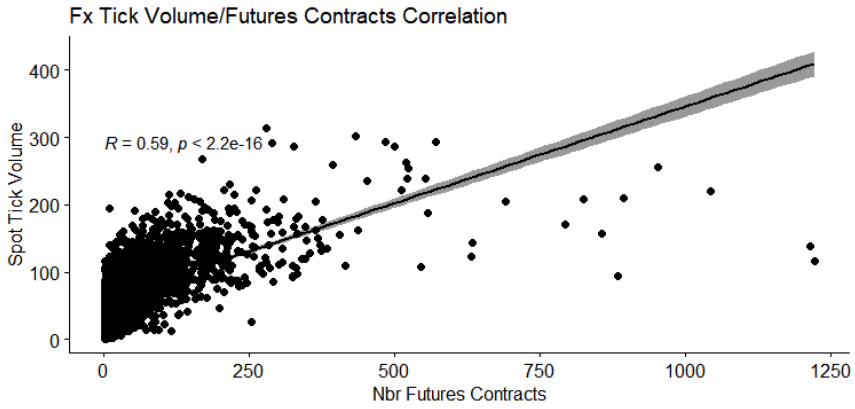


Fig 9. Fx futures volume and Fx tick volume show only a moderate degree of correlation

Summary

The relatively high correlation of intraday FX spot tick volumes from a number of different brokers /market makers supports the contention, intimated by Marney some years ago that tick volume is highly correlated to "real" market volume. To the best of my knowledge, this theory has never been seriously challenged or dis-proven although there are many who chose to blissfully ignore it and maintain support for the idea that tick volume is useless to traders.

Based on the above assumption, the moderate correlation of FX futures volume with tick volume supports the idea that FX futures data can be used as a reasonable but sub optimal proxy for the spot FX market. FX futures data has the advantage that it can provide Depth Of Market (DOM) facilities that can be used to help predict price moves. A disadvantage of FX futures data is that is quite expensive. There are typically set up costs and ongoing fees. Traders with only a moderate amount of capital and/or a low risk appetite may find the costs prohibitive.

It needs to be pointed out that these tests were carried out using only one currency pair (GBPUSD) and only two days of trading data. It would be most unwise to conclude that these tests provide conclusive evidence. They are indicative only. It would seem worthwhile for others to follow up this inquiry with further testing of other currency pairs over longer periods of time. The R scripts used for this evaluation are freely available and may be useful for this purpose.

Testing of this nature provides a much more objective approach to trading than the common practise of just blindly following someone else's subjective and untested opinions.

The creative use of tick volume in trading appears to offer opportunities for traders to improve their edge. Such opportunities however are limited to those who are able to overcome their inculcated beliefs about the uselessness of FX tick volume. Such beliefs can be powerful enough to prevent a person from being able to evaluate and utilise this sort of knowledge in a beneficial way.

References

Are price updates a good proxy for actual traded volume in FX?
By Caspar Marney, FX Trader Magazine
http://www.fxtradermagazine.com/pdf/previewSpecialEdition1.pdf

Converting Forex Volume Into Dollars - R File and data.zip
Available from the authors web site www.MindYourTrading.com .Please use the contacts page to request this file if you don't have access to the downloads section.

Credits
Thanks to PeterCalab, RickM, Pumi and giselle1723 from the Trading Discussion thread on the Forex Factory forum, "The brutal truth of day trading" for sharing their ideas, insights and FX futures data.