NINJATRADER®

TECHNICAL ANALYSIS OF STOCKS & COMMODITIES
READERS' CHOICE AWARD

We are honored to be recognized as a top trading software provider in the 2020 Readers' Choice Awards.

WINNER: Trading Software > $1000
WINNER: Trading Software < $500

Thank you to all who voted for NinjaTrader's industry leading platform & brokerage services.

ninjatrader.com

IBKR charges margin loan rates[1] from 0.75% to 1.59%

Start Investing Today!

ibkr.com/iwantmore

Interactive Brokers **Rated #1**
Best Online Broker 2020 by Barron's*

InteractiveBrokers

Trading on margin is **only for sophisticated investors** with **high risk tolerance**. You may lose more than your initial investment.

Member - NYSE, FINRA, SIPC – *Interactive Brokers rated #1, Best Online Broker according to Barron's Best Online Brokers Survey of 2020: February 21, 2020. For more information see, ibkr.com/info - Barron's is a registered trademark of Dow Jones & Co. Inc. [1] Margin Loan rate as of 07/08/2020. IB calculates the interest charged on margin loans using the applicable rates for each interest rate tier listed on its website. Rates shown apply to IBKR Pro clients only. Rates subject to change. For additional information on margin loan rates, see ibkr.com/interest

10-IB20-1374CH1358

BRUCE WALDMAN

Liberté! Fraternité!

How Good Is Your Data?

All data is equal — at least that's what we think.

by Sunny J. Harris

Is all data equal? If truth be told, I never gave it much thought. I have been using one vendor nearly exclusively for about 20 years. My fills are good enough. My closing prices seem to match what I see on television or find online. As long as the profits roll in, there has been no reason to question the data.

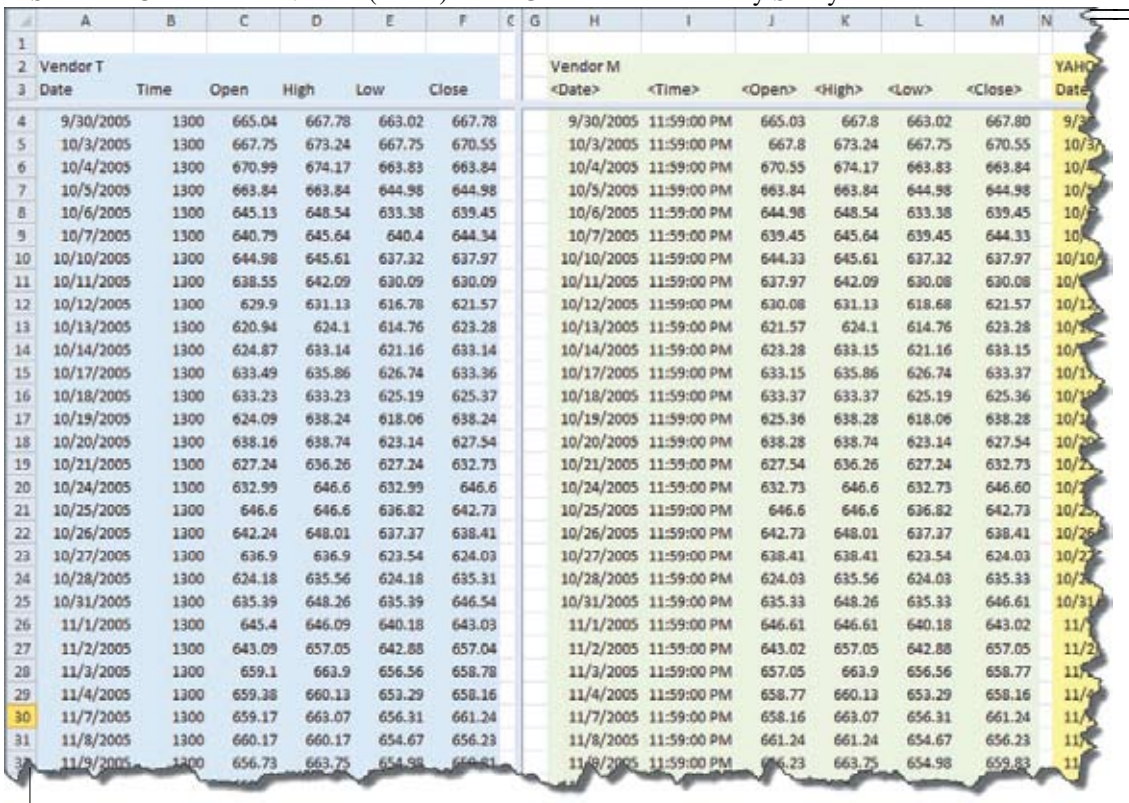
But then I was told by another vendor that my vendor's data is off by just enough to generate a side income, through the slippage from actual price to the price I am presented. My curiosity was piqued, and so I decided to investigate. First, I set up a spreadsheet and compared the two vendors. To keep it simple, I considered only the past five years of data. My data experiment ran from June 30, 2005, to June 29, 2010.

Running Experiments

I began by exporting the data for a single symbol from each software application to a comma-separated value (CSV) text file. The instrument I chose was the Russell 2000 index, which has different symbols in different software, like RUT, $RUT, and RU2000. I selected the Russell 2000 because of its high liquidity, ease of use, and it is something little guys like us can trade.

Figure 1 shows the beginning of the spreadsheet, with the data of the two vendors (T and M) in the columns. At first glance it appeared that everything was in order, with small discrepancies here and there. The differences in the data, where there is one, seem to be out in the hundredths place, like 600.01 vs. 600.02. That wouldn't make much difference over time, with some errors to the positive and the negative. It seems like it should be a wash.

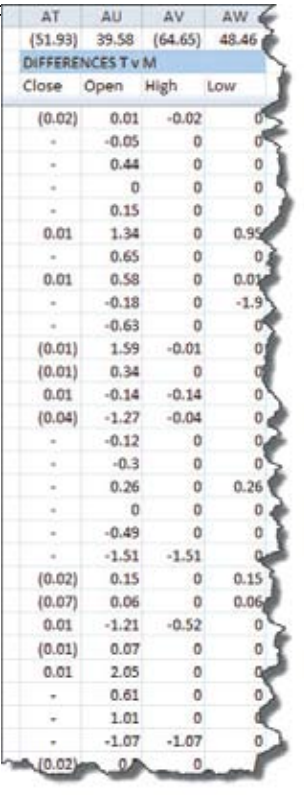
Next, I put columns in the spreadsheet to calculate the differences between the open, high, low, and close (OHLC) of each



	A	B	C	D	E	F
1						
2	Vendor T					
3	Date	Time	Open	High	Low	Close
4	9/30/2005	1300	665.04	667.78	663.02	667.78
5	10/3/2005	1300	667.75	673.24	667.75	670.55
6	10/4/2005	1300	670.99	674.17	663.83	663.84
7	10/5/2005	1300	663.84	663.84	644.98	644.98
8	10/6/2005	1300	645.13	648.54	633.38	639.45
9	10/7/2005	1300	640.79	645.64	640.4	644.34
10	10/10/2005	1300	644.98	645.61	637.32	637.97
11	10/11/2005	1300	638.55	642.09	630.09	630.09
12	10/12/2005	1300	629.9	631.13	616.78	621.57
13	10/13/2005	1300	620.94	624.1	614.76	623.28
14	10/14/2005	1300	624.87	633.14	621.16	633.14
15	10/17/2005	1300	633.49	635.86	626.74	633.36
16	10/18/2005	1300	633.23	633.23	625.19	625.37
17	10/19/2005	1300	624.09	638.24	618.06	638.24
18	10/20/2005	1300	638.16	638.74	623.14	627.54
19	10/21/2005	1300	627.24	636.26	627.24	632.73
20	10/24/2005	1300	632.99	646.6	632.99	646.6
21	10/25/2005	1300	646.6	646.6	636.82	642.73
22	10/26/2005	1300	642.24	648.01	637.37	638.41
23	10/27/2005	1300	636.9	636.9	623.54	624.03
24	10/28/2005	1300	624.18	635.56	624.18	635.31
25	10/31/2005	1300	635.39	648.26	635.39	646.54
26	11/1/2005	1300	645.4	646.09	640.18	643.03
27	11/2/2005	1300	643.09	657.05	642.88	657.04
28	11/3/2005	1300	659.1	663.9	656.56	658.78
29	11/4/2005	1300	659.38	660.13	653.29	658.16
30	11/7/2005	1300	659.17	663.07	656.31	661.24
31	11/8/2005	1300	660.17	660.17	654.67	656.23
32	11/9/2005	1300	656.73	663.75	654.98	659.83

	H	I	J	K	L	M	N
1							
2	Vendor M						YAHO
3	<Date>	<Time>	<Open>	<High>	<Low>	<Close>	Date
4	9/30/2005	11:59:00 PM	665.03	667.8	663.02	667.80	9/3
5	10/3/2005	11:59:00 PM	667.8	673.24	667.75	670.55	10/3
6	10/4/2005	11:59:00 PM	670.55	674.17	663.83	663.84	10/4
7	10/5/2005	11:59:00 PM	663.84	663.84	644.98	644.98	10/5
8	10/6/2005	11:59:00 PM	644.98	648.54	633.38	639.45	10/6
9	10/7/2005	11:59:00 PM	639.45	645.64	639.45	644.33	10/7
10	10/10/2005	11:59:00 PM	644.33	645.61	637.32	637.97	10/10
11	10/11/2005	11:59:00 PM	637.97	642.09	630.08	630.08	10/11
12	10/12/2005	11:59:00 PM	630.08	631.13	618.68	621.57	10/12
13	10/13/2005	11:59:00 PM	621.57	624.1	614.76	623.28	10/13
14	10/14/2005	11:59:00 PM	623.28	633.15	621.16	633.15	10/14
15	10/17/2005	11:59:00 PM	633.15	635.86	626.74	633.37	10/17
16	10/18/2005	11:59:00 PM	633.37	633.37	625.19	625.36	10/18
17	10/19/2005	11:59:00 PM	625.36	638.28	618.06	638.28	10/19
18	10/20/2005	11:59:00 PM	638.28	638.74	623.14	627.54	10/20
19	10/21/2005	11:59:00 PM	627.54	636.26	627.24	632.73	10/21
20	10/24/2005	11:59:00 PM	632.73	646.6	632.73	646.60	10/24
21	10/25/2005	11:59:00 PM	646.6	646.6	636.82	642.73	10/25
22	10/26/2005	11:59:00 PM	642.73	648.01	637.37	638.41	10/26
23	10/27/2005	11:59:00 PM	638.41	638.41	623.54	624.03	10/27
24	10/28/2005	11:59:00 PM	624.03	635.56	624.03	635.33	10/28
25	10/31/2005	11:59:00 PM	635.33	648.26	635.33	646.61	10/31
26	11/1/2005	11:59:00 PM	646.61	646.61	640.18	643.02	11/1
27	11/2/2005	11:59:00 PM	643.02	657.05	642.88	657.05	11/2
28	11/3/2005	11:59:00 PM	657.05	663.9	656.56	658.77	11/3
29	11/4/2005	11:59:00 PM	658.77	660.13	653.29	658.16	11/4
30	11/7/2005	11:59:00 PM	658.16	663.07	656.31	661.24	11/7
31	11/8/2005	11:59:00 PM	661.24	661.24	654.67	656.23	11/8
32	11/9/2005	11:59:00 PM	656.23	663.75	654.98	659.83	11/9

FIGURE 1: DATA COMPARISON, VENDOR T VS. VENDOR M. The differences in the data seems to be in the hundredths. Will deviations this small affect your profits?



	AT	AU	AV	AW
1	(51.93)	39.58	(64.65)	48.46
2	DIFFERENCES T v M			
3	Close	Open	High	Low
4	(0.02)	0.01	-0.02	0
5	-	-0.05	0	0
6	-	0.44	0	0
7	-	0	0	0
8	-	0.15	0	0
9	0.01	1.34	0	0.95
10	-	0.65	0	0
11	0.01	0.58	0	0.01
12	-	-0.18	0	-1.9
13	-	-0.63	0	0
14	(0.01)	1.59	-0.01	0
15	(0.01)	0.34	0	0
16	0.01	-0.14	-0.14	0
17	(0.04)	-1.27	-0.04	0
18	-	-0.12	0	0
19	-	-0.3	0	0
20	-	0.26	0	0.26
21	-	0	0	0
22	-	-0.49	0	0
23	-	-1.51	-1.51	0
24	(0.02)	0.15	0	0.15
25	(0.07)	0.06	0	0.06
26	0.01	-1.21	-0.52	0
27	(0.01)	0.07	0	0
28	0.01	2.05	0	0
29	-	0.61	0	0
30	-	1.01	0	0
31	-	-1.07	-1.07	0
32	(0.02)	0	0	

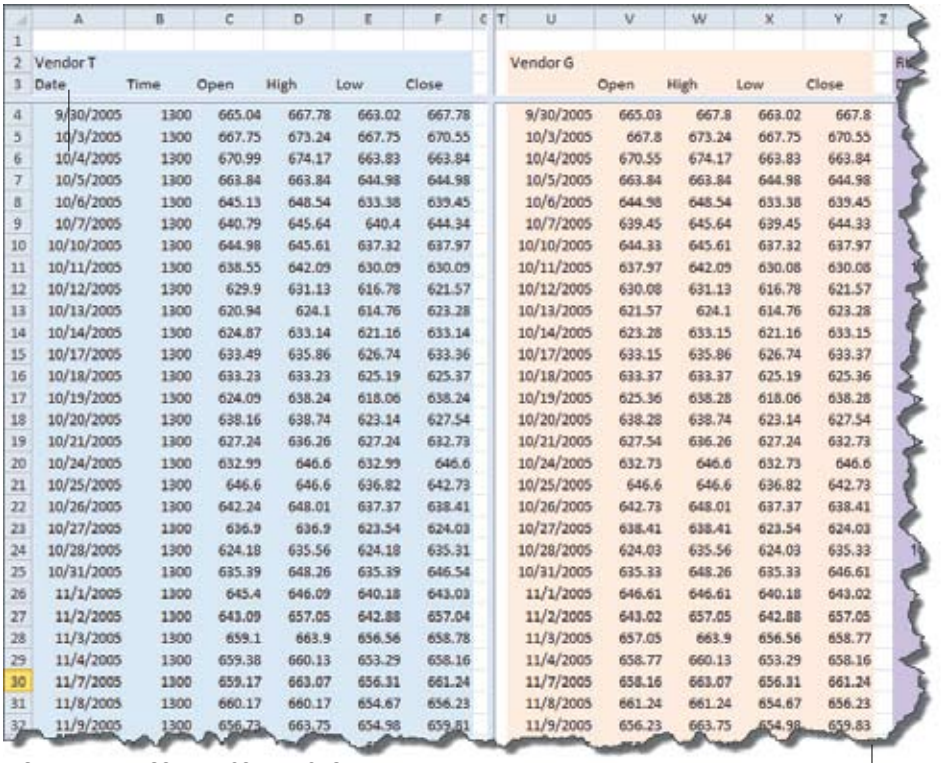
FIGURE 2: TOTAL DIFFERENCES, T VS. M. In the first row of data you see a sum of the differences in the open, high, low, and close between the two vendors. Now it gets interesting. The closes are 52 points lower, the opens are 40 points higher, the highs are 65 points lower, and the lows are 48 points higher.

vendor. Part of that spreadsheet is shown in Figure 2. At the top of each column, in the first row of data, is the result of calculating the sum of all the differences between the two vendors' OHLC data. I wouldn't have been surprised if each component had been consistently lower or higher than the other. But these summation numbers show that the data is all over the map. The closes are 52 points lower, the opens are 40 points higher, the highs are 65 points lower, and the lows are 48 points higher. The spread between the numbers is alternating positive and negative. Could it be — as one vendor suggested — that there is enough of a spread in there for vendor T to cash in on the spread alone?

With this information, I wanted to compare the data I had come up with to another well-known vendor to see whether their data matched either vendor T or vendor M. So I went back to the Export facility in software G to create another set of columns in the spreadsheet. I hoped that the data from vendor G would match one or the other of the first two vendors and I would come up with an answer.

Figure 3 shows the data from vendor T and vendor G, the new set for comparison. Again, I gave the data a cursory glance, but nothing seemed amiss. The variations are again out in the hundredths place.

You can do all the testing in the world, but when it comes to entering real trades, the markets will hand you something unexpected.



	A	B	C	D	E	F
1						
2	Vendor T					
3	Date	Time	Open	High	Low	Close
4	9/30/2005	1300	665.04	667.78	663.02	667.78
5	10/3/2005	1300	667.75	673.24	667.75	670.55
6	10/4/2005	1300	670.99	674.17	663.83	663.84
7	10/5/2005	1300	663.84	663.84	644.98	644.98
8	10/6/2005	1300	645.13	648.54	633.38	639.45
9	10/7/2005	1300	640.79	645.64	640.4	644.34
10	10/10/2005	1300	644.98	645.61	637.32	637.97
11	10/11/2005	1300	638.55	642.09	630.09	630.09
12	10/12/2005	1300	629.9	631.13	616.78	621.57
13	10/13/2005	1300	620.94	624.1	614.76	623.28
14	10/14/2005	1300	624.87	633.14	621.16	633.14
15	10/17/2005	1300	633.49	635.86	626.74	633.36
16	10/18/2005	1300	633.23	633.23	625.19	625.37
17	10/19/2005	1300	624.09	638.24	618.06	638.24
18	10/20/2005	1300	638.16	638.74	623.14	627.54
19	10/21/2005	1300	627.24	636.26	627.24	632.73
20	10/24/2005	1300	632.99	646.6	632.99	646.6
21	10/25/2005	1300	646.6	646.6	636.82	642.73
22	10/26/2005	1300	642.24	648.01	637.37	638.41
23	10/27/2005	1300	636.9	636.9	623.54	624.03
24	10/28/2005	1300	624.18	635.56	624.18	635.31
25	10/31/2005	1300	635.39	648.26	635.39	646.54
26	11/1/2005	1300	645.4	646.09	640.18	643.03
27	11/2/2005	1300	643.09	657.05	642.88	657.04
28	11/3/2005	1300	659.1	663.9	656.56	658.78
29	11/4/2005	1300	659.38	660.13	653.29	658.16
30	11/7/2005	1300	659.17	663.07	656.31	661.24
31	11/8/2005	1300	660.17	660.17	654.67	656.23
32	11/9/2005	1300	656.73	663.75	654.98	659.83

	U	V	W	X	Y	Z
1						
2	Vendor G					
3		Open	High	Low	Close	
4	9/30/2005	665.03	667.8	663.02	667.8	
5	10/3/2005	667.8	673.24	667.75	670.55	
6	10/4/2005	670.55	674.17	663.83	663.84	
7	10/5/2005	663.84	663.84	644.98	644.98	
8	10/6/2005	644.98	648.54	633.38	639.45	
9	10/7/2005	639.45	645.64	639.45	644.33	
10	10/10/2005	644.33	645.61	637.32	637.97	
11	10/11/2005	637.97	642.09	630.08	630.08	
12	10/12/2005	630.08	631.13	616.78	621.57	
13	10/13/2005	621.57	624.1	614.76	623.28	
14	10/14/2005	623.28	633.15	621.16	633.15	
15	10/17/2005	633.15	635.86	626.74	633.37	
16	10/18/2005	633.37	633.37	625.19	625.36	
17	10/19/2005	625.36	638.28	618.06	638.28	
18	10/20/2005	638.28	638.74	623.14	627.54	
19	10/21/2005	627.54	636.26	627.24	632.73	
20	10/24/2005	632.73	646.6	632.73	646.6	
21	10/25/2005	646.6	646.6	636.82	642.73	
22	10/26/2005	642.73	648.01	637.37	638.41	
23	10/27/2005	638.41	638.41	623.54	624.03	
24	10/28/2005	624.03	635.56	624.03	635.33	
25	10/31/2005	635.33	648.26	635.33	646.61	
26	11/1/2005	646.61	646.61	640.18	643.02	
27	11/2/2005	643.02	657.05	642.88	657.05	
28	11/3/2005	657.05	663.9	656.56	658.77	
29	11/4/2005	658.77	660.13	653.29	658.16	
30	11/7/2005	658.16	663.07	656.31	661.24	
31	11/8/2005	661.24	661.24	654.67	656.23	
32	11/9/2005	656.23	663.75	654.98	659.83	

FIGURE 3: DATA COMPARISON, T VS. G. The variations are, for the most part, in the hundredths. Nothing seems amiss.

AY	AZ	BA	BB	BC
(61.10)	26.61	(64.37)	47.27	
DIFFERENCES T v G				
Close	Open	High	Low	
-0.02	0.01	-0.02	0	
0	-0.05	0	0	
0	0.44	0	0	
0	0	0	0	
0	0.15	0	0	
0.01	1.34	0	0.95	
0	0.65	0	0	
0.01	0.58	0	0.01	
0	-0.18	0	0	
0	-0.63	0	0	
-0.01	1.59	-0.01	0	
-0.01	0.34	0	0	
0.01	-0.14	-0.14	0	
-0.04	-1.27	-0.04	0	
0	-0.12	0	0	
0	-0.3	0	0	
0	0.26	0	0.26	
0	0	0	0	
0	-0.49	0	0	
0	-1.51	-1.51	0	
-0.02	0.15	0	0.15	
-0.07	0.06	0	0.06	
0.01	-1.21	-0.52	0	
-0.01	0.07	0	0	
0.01	2.05	0	0	
0	0.61	0	0	
0	1.01	0	0	
0	-1.07	-1.07	0	
-0.02	0.5	0	0	

FIGURE 4: TOTAL DIFFERENCES, T VS. G. Although the discrepancies are small, when summed over five years of data it could accumulate to $12.57 per point on the Russell 2000.

I found small discrepancies that led to large numbers when summed over time (Figure 4). On its own, an error of 0.01 doesn't seem like much. But when you add that up over five years of data, it is 1,257 trading days and an accumulated error of $12.57. Remember, each point is worth $100 on the RUT.

This is where it starts looking scary. Multiplying $12.57 x 100 gives you $1,257. That's over $1,000 out of the trader's pocket. It isn't huge, but if you are the vendor and you have 20,000 clients at $1,000 each, that comes to $20 million. That is $20 million over five years. *Now* I was beginning to understand what that vendor was talking about.

Still, I couldn't go anywhere with this bit of information. This situation was akin to having a clock shop where each clock tells time a bit off from every other clock in the shop. There's no way to tell what time it really is. Which clock is telling the right time?

This situation demands that I compare the data from vendor G to vendor T and also to vendor M. I'm not sure what I would find out if none of it matched, but if one matched one other, then I'll know something about the veracity of the vendor's data that didn't match.

Here's the spreadsheet I have for three vendors' data so far (see Figure 5). Back to the differences spreadsheet, I inserted columns for calculating the new spreads: vendors T versus G; T versus M; and G versus M. That setup will be compared against the other and maybe I'll get some clarity. The differences section of the spreadsheet can be seen in Figure 6.

Aha! Look at the zeroes in columns BD through BG. Reading the description in row B over those columns (shaded green), I see that the zeroes show up when comparing vendors M to G. Still, looking at the numbers over the header "Differences M v G," we see that despite all the zeroes there are discrepancies along the way, giving us (9.17) among the closing values.

As I scanned the columns of this spreadsheet comparison, I found that on September 17, 2008, there was a difference of (8.83) between the close of M and the close of G. That was where most of the error comes in.

How could these vendors have such differences among their data? Isn't the close the close, no matter who vends it?

Next, I called the Russell 2000 exchange and got the data

	A	B	C	D	E	F	G	H	I	J	K	L	R	S	T	U	V	W	X
1																			
2	Vendor T						Vendor M							Vendor G					RUSSEL
3	Date	Open	High	Low	Close		<Date>	<Open>	<High>	<Low>	<Close>				Open	High	Low	Close	Date
4	9/30/2005	665.04	667.78	663.02	667.78		9/30/2005	665.03	667.8	663.02	667.80			9/30/2005	665.03	667.8	663.02	667.8	9/3
5	10/3/2005	667.75	673.24	667.75	670.55		10/3/2005	667.8	673.24	667.75	670.55			10/3/2005	667.8	673.24	667.75	670.55	10
6	10/4/2005	670.99	674.17	663.83	663.84		10/4/2005	670.55	674.17	663.83	663.84			10/4/2005	670.55	674.17	663.83	663.84	10/4
7	10/5/2005	663.84	663.84	644.98	644.98		10/5/2005	663.84	663.84	644.98	644.98			10/5/2005	663.84	663.84	644.98	644.98	10/5
8	10/6/2005	645.13	648.54	633.38	639.45		10/6/2005	644.98	648.54	633.38	639.45			10/6/2005	644.98	648.54	633.38	639.45	10
9	10/7/2005	640.79	645.64	640.4	644.34		10/7/2005	639.45	645.64	639.45	644.33			10/7/2005	639.45	645.64	639.45	644.33	10
10	10/10/2005	644.98	645.61	637.32	637.97		10/10/2005	644.33	645.61	637.32	637.97			10/10/2005	644.33	645.61	637.32	637.97	10/1
11	10/11/2005	638.55	642.09	630.09	630.09		10/11/2005	637.97	642.09	630.08	630.08			10/11/2005	637.97	642.09	630.08	630.08	10/1
12	10/12/2005	629.9	631.13	616.78	621.57		10/12/2005	630.08	631.13	618.68	621.57			10/12/2005	630.08	631.13	616.78	621.57	10/1
13	10/13/2005	620.94	624.1	614.76	623.28		10/13/2005	621.57	624.1	614.76	623.28			10/13/2005	621.57	624.1	614.76	623.28	10/1
14	10/14/2005	624.87	633.14	621.16	633.14		10/14/2005	623.28	633.15	621.16	633.15			10/14/2005	623.28	633.15	621.16	633.15	10/14
15	10/17/2005	633.49	635.86	626.74	633.36		10/17/2005	633.15	635.86	626.74	633.37			10/17/2005	633.15	635.86	626.74	633.37	10/1
16	10/18/2005	633.23	633.23	625.19	625.37		10/18/2005	633.37	633.37	625.19	625.36			10/18/2005	633.37	633.37	625.19	625.36	10/
17	10/19/2005	624.09	638.24	618.06	638.24		10/19/2005	625.36	638.28	618.06	638.28			10/19/2005	625.36	638.28	618.06	638.28	10/1
18	10/20/2005	638.16	638.74	623.14	627.54		10/20/2005	638.28	638.74	623.14	627.54			10/20/2005	638.28	638.74	623.14	627.54	10/2
19	10/21/2005	627.24	636.26	627.24	632.73		10/21/2005	627.54	636.26	627.24	632.73			10/21/2005	627.54	636.26	627.24	632.73	10/
20	10/24/2005	632.99	646.6	632.99	646.6		10/24/2005	632.73	646.6	632.73	646.60			10/24/2005	632.73	646.6	632.73	646.6	10/24
21	10/25/2005	646.6	646.6	636.82	642.73		10/25/2005	646.6	646.6	636.82	642.73			10/25/2005	646.6	646.6	636.82	642.73	10/25
22	10/26/2005	642.24	648.01	637.37	638.41		10/26/2005	642.73	648.01	637.37	638.41			10/26/2005	642.73	648.01	637.37	638.41	10/2
23	10/27/2005	636.9	636.9	623.54	624.03		10/27/2005	638.41	638.41	623.54	624.03			10/27/2005	638.41	638.41	623.54	624.03	10/2
24	10/28/2005	624.18	635.56	624.18	635.31		10/28/2005	624.03	635.56	624.03	635.33			10/28/2005	624.03	635.56	624.03	635.33	10/28
25	10/31/2005	635.39	648.26	635.39	646.54		10/31/2005	635.33	648.26	635.33	646.61			10/31/2005	635.33	648.26	635.33	646.61	10/31
26	11/1/2005	645.4	646.09	640.18	643.03		11/1/2005	646.61	646.61	640.18	643.02			11/1/2005	646.61	646.61	640.18	643.02	11/1
27	11/2/2005	643.09	657.05	642.88	657.04		11/2/2005	643.02	657.05	642.88	657.05			11/2/2005	643.02	657.05	642.88	657.05	11/2
28	11/3/2005	659.1	663.9	656.56	658.78		11/3/2005	657.05	663.9	656.56	658.77			11/3/2005	657.05	663.9	656.56	658.77	11/
29	11/4/2005	659.38	660.13	653.29	658.16		11/4/2005	658.77	660.13	653.29	658.16			11/4/2005	658.77	660.13	653.29	658.16	11/
30	11/7/2005	659.17	663.07	656.31	661.24		11/7/2005	658.16	663.07	656.31	661.24			11/7/2005	658.16	663.07	656.31	661.24	11
31	11/8/2005	660.17	660.17	654.67	656.23		11/8/2005	661.24	661.24	654.67	656.23			11/8/2005	661.24	661.24	654.67	656.23	11
32	11/9/2005	656.23	663.75	654.98	659.83		11/9/2005	656.23	663.75	654.98	659.83			11/9/2005	656.23	663.75	654.98	659.83	1

FIGURE 5: THREE VENDORS' DATA. Comparing data from three vendors will say something about the veracity of the data.

AT	AU	AV	AW	AX	AY	AZ	BA	BB	BC	BD	BE	BF	BG
(51.93)	39.58	(64.65)	48.46		(61.10)	26.61	(64.37)	47.27		(9.17)	(12.97)	0.28	(1.19)
DIFFERENCES T v M					DIFFERENCES T v G					DIFFERENCES M v G			
Close	Open	High	Low		Close	Open	High	Low		Close	Open	High	Low
(0.02)	0.01	-0.02	0		-0.02	0.01	-0.02	0		-	0	0	0
-	-0.05	0	0		0	-0.05	0	0		-	0	0	0
-	0.44	0	0		0	0.44	0	0		-	0	0	0
-	0	0	0		0	0	0	0		-	0	0	0
-	0.15	0	0		0	0.15	0	0		-	0	0	0
0.01	1.34	0	0.95		0.01	1.34	0	0.95		-	0	0	0
-	0.65	0	0		0	0.65	0	0		-	0	0	0
0.01	0.58	0	0.01		0.01	0.58	0	0.01		-	0	0	0
-	-0.18	0	-1.9		0	-0.18	0	0		-	0	0	1.9
-	-0.63	0	0		0	-0.63	0	0		-	0	0	0
(0.01)	1.59	-0.01	0		-0.01	1.59	-0.01	0		-	0	0	0
(0.01)	0.34	0	0		-0.01	0.34	0	0		-	0	0	0
0.01	-0.14	-0.14	0		0.01	-0.14	-0.14	0		-	0	0	0
(0.04)	-1.27	-0.04	0		-0.04	-1.27	-0.04	0		-	0	0	0
-	-0.12	0	0		0	-0.12	0	0		-	0	0	0
-	-0.3	0	0		0	-0.3	0	0		-	0	0	0
-	0.26	0	0.26		0	0.26	0	0.26		-	0	0	0
-	0	0	0		0	0	0	0		-	0	0	0
-	-0.49	0	0		0	-0.49	0	0		-	0	0	0
-	-1.51	-1.51	0		0	-1.51	-1.51	0		-	0	0	0
(0.02)	0.15	0	0.15		-0.02	0.15	0	0.15		-	0	0	0
(0.07)	0.06	0	0.06		-0.07	0.06	0	0.06		-	0	0	0
0.01	-1.21	-0.52	0		0.01	-1.21	-0.52	0		-	0	0	0
(0.01)	0.07	0	0		-0.01	0.07	0	0		-	0	0	0
0.01	2.05	0	0		0.01	2.05	0	0		-	0	0	0
-	0.61	0	0		0	0.61	0	0		-	0	0	0
-	1.01	0	0		0	1.01	0	0		-	0	0	0
-	-1.07	-1.07	0		0	-1.07	-1.07	0		-	0	0	0
(0.02)	0.5	0	0		-0.02	0.5	0	0		-	0	0	0

FIGURE 6: TOTAL DIFFERENCES, T VS. M VS. G. Even though there appears to be hardly any discrepancies between M and G, the totals reflect another story.

from them. Now that data should be correct, right? It's their own index, so they should know. But they only had closing prices for the first part of my experiment (September 30, 2005, through June, 22, 2007). Closes will have to do for comparing data vendors versus the Russell 2000 itself. Their closing prices are accurate down to six decimal places, while the others only have two digits after the decimal.

Adding yet another set of columns to the spreadsheet, I placed the RU2000 from the exchange in place. While I was at it, I introduced another column with the calculation for the range of each day. As I looked at the data, I wondered whether the errors worked themselves out by having the same range for the day, even if the open and close were different. That's why there's a column labeled "Range." Figure 7 shows three data vendors and the Russell 2000 exchange data side by side.

The more columns I added, the more difficult it is for you to read. So for those interested in the details found in these spreadsheets, visit www.MoneyMentor.com/Articles.html, where you can see enlargements of these figures.

When you compare the closes of the Russell index to the closes of vendors G, T, or M, there are slight discrepancies. I decided to add another data vendor, one who does not connect a brokerage firm to the data (as far as I know). Yahoo! makes its data available for free, and because it is such a popular data source, it should have pretty clean data.

As I looked closer at the data, I saw that some vendors were using the first opened trade for the open value, and some

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	V	W	X	Y	Z	AA	AB	AC	AD
1																							
2	Vendor T							Vendor M							Vendor G							RUSSELL EXCHANGE	
3	Date	Open	High	Low	Close	Range		<Date>	<Open>	<High>	<Low>	<Close>	Range		Date	Open	High	Low	Close	Range		Date	Close
4	9/30/2005	665.04	667.78	663.02	667.78	4.76		9/30/2005	665.03	667.8	663.02	667.80	4.78		9/30/2005	665.03	667.8	663.02	667.8	4.78		9/30/2005	667.798000
5	10/3/2005	667.75	673.24	667.75	670.55	5.49		10/3/2005	667.8	673.24	667.75	670.55	5.49		10/3/2005	667.8	673.24	667.75	670.55	5.49		10/3/2005	670.549600
6	10/4/2005	670.99	674.17	663.83	663.84	10.34		10/4/2005	670.55	674.17	663.83	663.84	10.34		10/4/2005	670.55	674.17	663.83	663.84	10.34		10/4/2005	663.839800
7	10/5/2005	663.84	663.84	644.98	644.98	18.86		10/5/2005	663.84	663.84	644.98	644.98	18.86		10/5/2005	663.84	663.84	644.98	644.98	18.86		10/5/2005	644.976700
8	10/6/2005	645.13	648.54	633.38	639.45	15.16		10/6/2005	644.98	648.54	633.38	639.45	15.16		10/6/2005	644.98	648.54	633.38	639.45	15.16		10/6/2005	639.449900
9	10/7/2005	640.79	645.64	640.4	644.34	5.24		10/7/2005	639.45	645.64	639.45	644.33	6.19		10/7/2005	639.45	645.64	639.45	644.33	6.19		10/7/2005	644.334200
10	10/10/2005	644.98	645.61	637.32	637.97	8.29		10/10/2005	644.33	645.61	637.32	637.97	8.29		10/10/2005	644.33	645.61	637.32	637.97	8.29		10/10/2005	637.968800
11	10/11/2005	638.55	642.09	630.09	630.09	12		10/11/2005	637.97	642.09	630.08	630.08	12.01		10/11/2005	637.97	642.09	630.08	630.08	12.01		10/11/2005	630.083100
12	10/12/2005	629.9	631.13	616.78	621.57	14.35		10/12/2005	630.08	631.13	618.68	621.57	12.45		10/12/2005	630.08	631.13	616.78	621.57	14.35		10/12/2005	621.569300
13	10/13/2005	620.94	624.1	614.76	623.28	9.34		10/13/2005	621.57	624.1	614.76	623.28	9.34		10/13/2005	621.57	624.1	614.76	623.28	9.34		10/13/2005	623.280900
14	10/14/2005	624.87	633.14	621.16	633.14	11.98		10/14/2005	623.28	633.15	621.16	633.15	11.99		10/14/2005	623.28	633.15	621.16	633.15	11.99		10/14/2005	633.150700
15	10/17/2005	633.49	635.86	626.74	633.36	9.12		10/17/2005	633.15	635.86	626.74	633.37	9.12		10/17/2005	633.15	635.86	626.74	633.37	9.12		10/17/2005	633.369900
16	10/18/2005	633.23	633.23	625.19	625.37	8.04		10/18/2005	633.37	633.37	625.19	625.36	8.18		10/18/2005	633.37	633.37	625.19	625.36	8.18		10/18/2005	625.364500
17	10/19/2005	624.09	638.24	618.06	638.24	20.18		10/19/2005	625.36	638.28	618.06	638.28	20.22		10/19/2005	625.36	638.28	618.06	638.28	20.22		10/19/2005	638.281800
18	10/20/2005	638.16	638.74	623.14	627.54	15.6		10/20/2005	638.28	638.74	623.14	627.54	15.60		10/20/2005	638.28	638.74	623.14	627.54	15.6		10/20/2005	627.540600
19	10/21/2005	627.24	636.26	627.24	632.73	9.02		10/21/2005	627.54	636.26	627.24	632.73	9.02		10/21/2005	627.54	636.26	627.24	632.73	9.02		10/21/2005	632.732900
20	10/24/2005	632.99	646.6	632.99	646.6	13.61		10/24/2005	632.73	646.6	632.73	646.60	13.87		10/24/2005	632.73	646.6	632.73	646.6	13.87		10/24/2005	646.603500
21	10/25/2005	646.6	646.6	636.82	642.73	9.78		10/25/2005	646.6	646.6	636.82	642.73	9.78		10/25/2005	646.6	646.6	636.82	642.73	9.78		10/25/2005	642.732400
22	10/26/2005	642.24	648.01	637.37	638.41	10.64		10/26/2005	642.73	648.01	637.37	638.41	10.64		10/26/2005	642.73	648.01	637.37	638.41	10.64		10/26/2005	638.410100
23	10/27/2005	636.9	636.9	623.54	624.03	13.36		10/27/2005	638.41	638.41	623.54	624.03	14.87		10/27/2005	638.41	638.41	623.54	624.03	14.87		10/27/2005	624.028200
24	10/28/2005	624.18	635.56	624.18	635.31	11.38		10/28/2005	624.03	635.56	624.03	635.33	11.53		10/28/2005	624.03	635.56	624.03	635.33	11.53		10/28/2005	635.329100
25	10/31/2005	635.39	648.26	635.39	646.54	12.87		10/31/2005	635.33	648.26	635.33	646.61	12.93		10/31/2005	635.33	648.26	635.33	646.61	12.93		10/31/2005	646.605400
26	11/1/2005	645.4	646.09	640.18	643.03	5.91		11/1/2005	646.61	646.61	640.18	643.02	6.43		11/1/2005	646.61	646.61	640.18	643.02	6.43		11/1/2005	643.024100
27	11/2/2005	643.09	657.05	642.88	657.04	14.17		11/2/2005	643.02	657.05	642.88	657.05	14.17		11/2/2005	643.02	657.05	642.88	657.05	14.17		11/2/2005	657.051900
28	11/3/2005	659.1	663.9	656.56	658.78	7.34		11/3/2005	657.05	663.9	656.56	658.77	7.34		11/3/2005	657.05	663.9	656.56	658.77	7.34		11/3/2005	658.771800
29	11/4/2005	659.38	660.13	653.29	658.16	6.84		11/4/2005	658.77	660.13	653.29	658.16	6.84		11/4/2005	658.77	660.13	653.29	658.16	6.84		11/4/2005	658.162800
30	11/7/2005	659.17	663.07	656.31	661.24	6.76		11/7/2005	658.16	663.07	656.31	661.24	6.76		11/7/2005	658.16	663.07	656.31	661.24	6.76		11/7/2005	661.238500
31	11/8/2005	660.17	660.17	654.67	656.23	5.5		11/8/2005	661.24	661.24	654.67	656.23	6.57		11/8/2005	661.24	661.24	654.67	656.23	6.57		11/8/2005	656.234700
32	11/9/2005	656.73	663.75	654.98	659.81	8.77		11/9/2005	656.23	663.75	654.98	659.83	8.77		11/9/2005	656.23	663.75	654.98	659.83	8.77		11/9/2005	659.814300
33	11/10/2005	659.92	665.14	651.5	664.94	13.64		11/10/2005	659.81	665.14	651.5	664.93	13.64		11/10/2005	659.81	665.14	651.5	664.93	13.64		11/10/2005	664.934500
34	11/11/2005	664.22	667.82	663.43	666.66	4.39		11/11/2005	664.93	667.82	663.43	666.66	4.39		11/11/2005	664.93	667.82	663.43	666.66	4.39		11/11/2005	666.663900

FIGURE 7: RUSSELL 2000 EXCHANGE DATA VS. T VS. M VS. G. If you look at the data of the close carefully, you will note there are slight discrepancies.

FIGURE 8: TWO DATA VENDORS ON ONE CHART. Here you see two sets of data overlaid on top of each other. The orange tick is the data from the exchange and the blue and green ticks are from T. Some bars such as the first, third, and fifth do have a discrepancy.

were using the opening range of the first few minutes. The same applied to the close, in that some use the value at the bell and some use the range as all the last few orders trickle in. In Figure 8, I overlaid two sets of data. You can see where the orange tick is at a different location than the green tick. Orange stands for the Russell 2000 data from the exchange, while the underlying blue and green are from vendor T.

Looking at the chart provided a clearer picture. You can see right from the first bar on this chart that the open tick (to the left) has an orange one and a green one, only slightly different, but different nonetheless.

On the third bar from the left there is some difference between the two opens, though the closes are equal. Similarly on the fifth bar, you can visualize the discrepancy across the chart. The differences are subtle, but they are there.

As for my own trading, I entered and exited on market orders. Or I let a stop take me out. In neither case was it crucial that I placed orders on the open of the bar, even on the close of the bar.

However — and this is a *big* however — when writing and testing system ideas, many, if not most, coders specify things like:

```
IF condition1 THEN BUY next bar on the OPEN;
IF condition2 THEN SELL on CLOSE;
```

If I tested such code against these five datasets I would get different results, different profits and losses, depending on the data vendor or software vendor. So which results are correct? The correct data is the set that gives the same results as actual trades entered into the market would yield. And that brings us to the heart of the matter. Do you want to put your money into the markets in a reversal system like, say, the moving average convergence/divergence (MACD), just to see whether the trades it comes up with replicate the trades the hypothetical system generates? Of course not, and neither do I. So we are at an impasse.

I'll come back to the impasse in a minute, but for now let's get back to the data comparisons. From Figure 6 you can see that vendors M and G are very close in the data they provide. Most of the cells in the spreadsheet contain zeroes; there is no difference in the data point from vendor G and the data point from M. You can see a spreadsheet of the differences between all vendors on my website.

PUTTING IT ALL TOGETHER

It is clear that there are many differences between data vendors. The close is not the close all the world around, and there is no way to evaluate which is better. The better data is the data that most closely approximates what you would experience in actual trading scenarios. The problem is, I don't know how to run that experiment. I *could* set up an automated system in each software, where it would enter each trade in the markets on its own. Then, after letting the systems run for a year or so, we could compare the results of each trading experiment to evaluate the accuracy of the underlying data. Other than that, it is a matter of personal experience.

SYSTEM TESTING

The data between these five sources varies, sometimes widely. What if the data is different? It matters when you are entering trades in the markets, especially if you're trading at a very fast pace. It doesn't matter so much if you are off a penny in a trade that lasts for a year, or a month, or even a week. But if you are scalping for pennies, then the data you are making your decisions with needs to be exactly the same as the actual trades happening in the real market.

Running experiments down to the pennies is not within the scope of this article. I have limited the scope of these tests to daily charts over the past five years of data. This will illustrate the differences between the data sources when applied to hypothetical trading.

I will run the same experiment on all five sets of data. For vendors T, G, and M, the data is supplied by the vendor, so they go hand in hand. For vendor R, there is no software associated with the exchange, so I am going to import the data from a CSV text file into T's software and run the tests from there. The same applies to the data from Y: the data will be imported from a CSV file into T's software and tested from there. I will then view the results of the tests by looking at the performance reports correlated to each dataset.

In order to set up the experiment, it is necessary to hold constant as many variables as possible, so that you compare apples to apples and get meaningful results. Here are the constraints I employed:

- Trade only one contract
- Constrain the data to the time frame October 13, 2005, to October 13, 2010
- Do not allow pyramiding
- Limit the input values to 12, 26, 9

Stats	Vendor G	Vendor M	Vendor R	Vendor T	Vendor Y
Total net profit	<$221.00>	<$302.60>	<$6.17>	<$272.94>	<$275.03>
Profit factor	0.82	0.76	0.99	0.78	0.78
# Trades	102	99	55	100	100%
Profitable	30.4%	29.29%	34.55%	30.00%	30.00%
Avg trade NP	<$2.00>	<$3.06>	<$0.11>	<$2.73>	<$2.75>
Ratio avg win: Avg loss	1.89	1.83	1.88	1.83	1.82
Avg bars in trades	11.98	12.6	14.67	12.72	12.72
Account size req'd	$511.00	$508.80	$258.81	$507.80	$506.64
CPC Index	0.46	0.41	0.62	0.43	0.43

FIGURE 9: PERFORMANCE REPORTS FROM ALL FIVE VENDORS. Statistically, there is little difference among the vendors as far as performance data goes. However, note that R (the exchange) has the best performance overall.

- Enter at the market, not on the open or close of the signal bar

With these values in mind, I ran one test and compared the results. At the beginning, I got wildly different answers. Because of the different philosophies of the software vendors, it was challenging finding the locations of the settings of things like trading one contract versus trading 100 contracts at a time. But with diligence, I got them all set up identically.

If you were to look at the charts of the data from each vendor, from a visual perspective the results look similar. To inspect the data more closely, I broke it down into a tabular format. I am not going to display all of the statistics, only the most important to the analysis (Figure 9).

Statistically, there is little difference among the four vendors, as far as performance data goes. What stands out, however, is the difference between the collections of the vendors against the data from the exchange itself. Vendor R — the exchange — has the best performance overall and is the one dataset different from the others.

I'm not going to get into all the results of the experiments. For the purposes of this article, I am not trying to find whether the MACD system works, rather attempting to uncover discrepancies among the data available for analyzing and trading. Admittedly, this is one set of data on the Russell 2000, and one set of parameters for only the MACD reversal system. It is not a comprehensive test, and by any means not a full analysis. But it is useful for answering the question posed by one vendor when touting the accuracy of their data. The outcome is not dramatic. All the data vendors present a losing outcome for the standard MACD strategy.

The only one close to positive is the data from the exchange; the rest are all negative by the same amount. It would be interesting to run comprehensive tests over a variety of time frames, optimizing the parameters and using other types of orders besides just buying at the market. But it would be an extensive test with thousands of outcomes.

THE BOTTOM LINE

The trades entered into real-time markets will vary far more than the data for the trades in this simple experiment. Over approximately 100 trades, there is only a few dollars' variation. In real-life trading, your fills will have wider ranges as markets move faster or slower, and the range between bids and asks widen and narrow. In real-life trading, by trading just one share, you can easily lose double the amounts shown in just a few minutes, depending on how quickly your order is entered in a fast-moving market.

Backtesting is not meant to provide precise replication of what would happen in the real markets. It is meant to give you an overall impression of whether your concepts are viable. The markets never again do the exact same thing they did before. They may echo similar patterns, but they don't duplicate them precisely. You can do all the testing in the world, but when it comes down to entering real trades, the markets will hand you something unexpected.

Larry Williams said something that has always stuck with me: "It takes time to make profits." Making a few dollars at a time to make $100 a day is against my trading philosophy. I believe in mathematical analysis of patterns and detection of setups that are likely to predict important pivot points and turns in the markets.

In the end, all five of these data sources are perfectly adequate for testing and for trading. It's all relative.

A trader, author, computer programmer, and mathematician, Sunny Harris has been trading since 1981. The first printing of her first book, Trading 101: How To Trade Like A Pro, *sold out in two weeks, and continues to be a financial best-seller, and her second book,* Trading 102: Getting Down To Business, *also achieved record sales. In early 2000, Harris released* Electronic Day Trading 101, *followed by* Getting Started In Trading in 2001. *She may be contacted at MoneyMentor.com.*

RELATED READING

• www.MoneyMentor.com/Articles.html
‡Microsoft Excel ‡TradeStation

‡See Editorial Resource Index

"I am delighted to confidentially tell you our annual report this year is a veritable labyrinth that exhaustively obfuscates any of our credit default swaps transactions."

The trading platform you'd build (if you built trading platforms).

Customize your view of the market so you can focus on what's important to you with thinkorswim®. Create your own indicators with thinkScript®, visualize the market with custom charting, and use personal predictions to forecast the values of stocks you care about with the Company Profile tool. And, this award-winning investing experience is now commission-free.*

See what customization can do for your trades at **tdameritrade.com/thinkorswim**

TD Ameritrade® | **Where Smart Investors Get Smarter**℠

***Applies to U.S. exchange-listed stocks, ETFs, and options. A $0.65 per contract fee applies for options trades.**

Company Profile forecasts are hypothetical, based on user-defined assumptions, and not guaranteed. All investing involves risks, including loss of principal. TD Ameritrade, Inc., member FINRA/SIPC. © 2019 TD Ameritrade.

Winner Winner...

TECHNICAL ANALYSIS OF STOCKS & COMMODITIES READERS' CHOICE AWARD

MetaStock D/C • **Winner** • Trading Software $500 - $1000

Bollinger Band System • **Winner** • Plug-ins
Metastock R/T •**1st Runner Up** • Software over $1,000
XENITH • **Finalist** • Real-Time Market Data
Dr. Elders Trading Room • **Finalist** • Plug-ins
Bollinger Band Toolkit • **Honorable Mention** • Plug-ins

Get a FREE trial of MetaStock with Market Data!*
Visit MetaStock.com/TASC.

*restrictions apply. See metastock.com/tasc for details.

MetaStock

SUBSCRIBE OR RENEW TODAY!

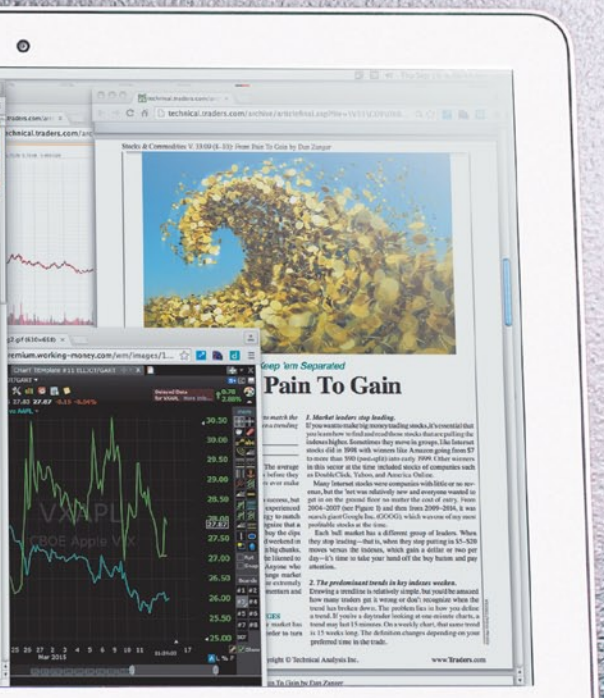
Every Stocks & Commodities subscription (regular and digital) includes:

- **Full access to our Digital Edition**
  The complete magazine as a PDF you can download.
- **Full access to our Digital Archives**
  That's 35 years' worth of content!
- **Complete access to WorkingMoney.com**
  The information you need to invest smartly and successfully.
- **Access to Traders.com Advantage**
  Insights, tips and techniques that can help you trade smarter.

1 year $89.99

2 years..... $149.99

3 years..... $199.99



PROFESSIONAL TRADERS' STARTER KIT

A 5-year subscription to S&C magazine that includes everything above PLUS a free* book, Charting The Stock Market: The Wyckoff Method, all for a price that saves you $150 off the year-by-year price! *Shipping & handling charges apply for foreign orders.

5 years..... $299.99

That's around $5 a month!

Visit **www.Traders.com** to find out more!

Email: Circ@Traders.com • Phone: 206-938-0570 facebook.com/STOCKSandCOMMODITIES @STOCKSandCOMM