ONLINE APPENDIX

Do Individual Currency Traders Make Money?

5.7 Robustness Checks with Second Data Set

The performance results from the main data set, presented in Panel B of Table 2, show that the top quartile of individual currency traders have alpha generating abilities, earning a statistically significant 0.59 percent per day (t-statistic=4.86). Additionally, our examination of trading characteristics shows that individual currency traders increase (decrease) trading based on the level of positive (negative) feedback received and outperform those who trade less frequently.

To test the robustness of the previous results we analyze a second data set which comprises account data from an additional online advisory service ZipSignals.com The robustness checks are performed due to possible sample selection bias concerns with our original date set and its relatively short sample time period (2004 - 2009). Analyzing a second data set from an alternative online advisory service, not only allows us to address the sample selection bias concern but also permits us to examine a more recent sample which addresses the inter-temporal stability of our results. A limitation of the new data set is that it contains only mean monthly return observations. Consequently, we cannot perform robustness checks of drawdown and market timing which are performed on the primary data set below. Furthermore, we are unable to test for the disposition effect which was performed on the Collective2 data above. Despite the limitations of the second data set it is important to analyze this data because it is difficult to obtain individual currency trader data from brokerages as they do not release individual currency trading data.

This new sample consists of 74 active accounts from July 2010 to August 2011. Accounts that have been closed or inactive were not provided in the database and are unavailable to these authors. The data include the account holder's name, the mean monthly gross return, the total number of trades, and the age of the account in weeks. To provide results comparable to those presented based on our primary data set we compute the mean daily gross return by dividing the mean monthly gross return by 20 (assuming 20 trading days per month) and calculate mean trades per day by dividing the total number of trades by the age of the account measured in days. Although the second data set does not contain transaction data, an analysis of the mean daily gross return provides insight into the performance of these traders. Furthermore, the mean number of trades per day allows us to test the calibration hypothesis, which predicts that individual currency traders who trade more frequently will outperform those who trade less frequently. One limitation of the secondary data set is that, unlike the Collective2 data which does not suffer from survivorship bias, account holders who close their accounts are not included in the data set. This creates survivorship bias. Consequently, it is quite possible that the performance and mean age may be higher in the secondary data set since underperformers are removed. We address these concerns with the analysis of the data below. We present descriptive statistics, performance results, and feedback trading results in Table 8.

Panel A of Table 8 reports descriptive statistics of gross returns, trade activity and the age of all 74 accounts. The mean daily gross return is 0.357 percent with the top quartile (bottom) quartile earning a gross return of 0.648 (0.005) percent per day, respectively. These results are similar to the results of our primary data set reported in Panel B of Table 2 which display cross-sectional variation in performance. Panel B of Table 2 reveals that the top performers (Q1) in our primary data set earn a gross return of 1.04 percent per day and the worst performers (Q4) earn a gross return of -0.25 percent per day. It is notable that the worst performers in the second data set (.005 percent per day) outperform the worst traders in the primary data set (-0.25 percent per day) by 0.255 percent per day. A likely explanation for this is survivorship bias since the secondary data set does not contain closed accounts. Two other noteworthy observations are mean trades per day and the age of accounts. The mean number of trades per day for the second data set is 2.35 and the mean age of accounts is 201.30 days. The mean number of trades shows that the individual currency traders in the secondary set are active traders but not as active as the traders in the primary data set where the mean trades per day is 3.31 (see Panel B of Table 1). A striking difference between the two data sets is the age of the accounts. As shown in Panel B of Table 1, the mean age for currency traders in the primary data set is 81.92 days. However, Panel A of Table 8 reports that the mean age for currency traders in the second data set is 201.30 days. The primary data set does not suffer from survivorship bias since it contains all accounts that were opened and closed during the sample time period. Consequently, all underperforming accounts are included in the database. Survivorship bias is present in the second database because accountholders who have closed their accounts are not present. A likely explanation for the age difference is that as traders experience losses, they close their accounts, and that tends to bias the results based on the second data set. Survivorship bias also occurs in currency hedge funds. Pojarliev and Levich (2010a) note that in their sample of currency hedge funds only 15 out of 22 funds survived over a three year period.

****Insert Table 8 about here****

We next examine the performance of the second data set by sorting accountholders into tertiles and report the results in Panel B of Table 8. Tertiles are used due to the number of observations. Quartile ranks provide similar results. The results presented in Panel B of Table 7 reveal that the top performing currency traders outperform the worst performing currency traders by 1.44 percent per day and it is significant (t-statistic=8.37). These results are similar to the primary data set in Table 2 where the difference between the top and worst performers is 1.29 percent per day and significant (t-statistic=8.63). Consequently, both data sets show that the top performers earn positive gross returns and the difference between the best and worst performers is significant.

Our final robustness check tests the feedback trading hypothesis which predicts a positive association between trade activity and performance. We test this by sorting the secondary data set by trade activity, proxied by mean trades per day. Accounts are ranked by mean trades per day and then divided into three groups. This is similar to the primary data set analysis performed in Table 5 where we report the most active traders in the first data set, proxied by mean trades per day, outperform the least active traders per day by 0.4359 percent per day and the difference is significant (t-statistic=3.43).

Panel C of Table 8 reports the results of sorts on trade activity for the second data set. The most active traders (T1) trade, on average, 4.873 times per day and earn a mean gross return of 0.628 per day. The least active traders (T3) trade, on average, 0.554 times per day and earn a gross return of 0.063 percent per day. The feedback hypothesis, which predicts that the difference in gross performance between the most active (T1) and least active traders (T3) will be positive is borne out in the data. Specifically, the most active traders outperform the least active by 0.57 percent per day and the difference is significant (t-statistic=3.45). This result is similar to all of the previous analyses performed on the primary data set which shows feedback can affect trading performance. Overall, the results of both data sets not only show that some individual currency traders are able to earn positive gross returns, but also there is a positive association between trade activity and gross performance.

5.8 Skill Measured by the Percentage of Winning Trades, Economic Significance of Profits, and Drawdowns

We next examine whether individual currency traders in this sample possess skill. To address this we first study individual transactions to determine whether their performance is determined by skill or by luck and then examine drawdown to find out whether individual currency traders possess skill at moderating losses.

5.8.1 Percentage of Winning Trades and Profit/Loss (P/L) Per Trade

Specifically, to gauge whether individual currency traders are skilled, we examine the percentage of winning trades for each account and then determine whether this percentage is statistically different from chance (a 50 percent win percentage). We start by counting the number of winning and losing trades for each account, where winning trades are defined as trades with a net profit greater than zero and losing trades are those with a net loss equal to or less than zero. The percentage of winning trades is the number of winning trades divided by the total number of trades per account. We then examine the full sample and performance-ranked quartiles (similar to our performance-based sorts above), with each quartile containing 107 accounts. The null hypothesis is that the percentage of winning trades will not be statistically different from chance (50 percent).

In addition to reporting the percentage of winning trades we also investigate the economic significance of profit and loss (P/L) per trade. This is performed because traders may have a high percentage of winning trades (> 50 percent) but the dollar loss of losing trades could be larger than dollar gain from winning trades, thus resulting in an economic loss. We analyze economic significance by calculating the mean P/L per individual trader then provide full-sample and cross-sectional result with ranks on performance similar to the percentage of winning trades discussed above. Table 9 presents the results of the analysis of individual trades.

****Insert Table 9 about here****

Panel A of Table 9 reports the full-sample results and shows that, on average, currency traders have winning trades 53.97 percent of the time, which is reliably different from 50 percent (t-statistic = 53.97). Furthermore, currency traders earn a mean $29.85 USD per trade and this is reliably different from zero (t-statistic=2.56). This provides support that the traders in this sample possess skill.

Panel B of Table 9 reports the results for the quartile performance sorts. Quartile 1, which contains the top-performing currency traders, shows that this group earns a profit on 66.78 percent of their trades, which is significantly different from 50 percent (t-statistic = 9.65). Additionally, they earn a mean profit of $248.31 per trade and this is also reliably different from zero (t-statistic=16.43). Quartile 2 traders earn a profit on 58.50 percent of their trades (t-statistic = 4.64) and $69.58 per trade (t-statistic=69.58). Quartile 3 traders have winning trades 48.33 percent of the time, which is not statistically different from 50 percent yet they lose -$66.71 per trade and this is significant (t-statistic=-2.26). Finally, the lowest-performing currency traders in quartile 4 are not skillful. Only 42.26 percent of their trades are profitable and they sustain a loss of -$601.21 (t-statistic=-17.42), on average, for each trade. Overall, the results of percentage of winning trades and the mean P/L per trade imply that approximately 50 percent of the individual currency traders analyzed in this sample posses skill trading spot currencies.

5.8.2 Drawdown Performance

Next we examine skill by investigating drawdown performance of individual currency traders. This analysis is expected to show the extent individual currency traders moderate their losses. A similar approach is used by Melvin and Shand (2011) to assess the skill of professional currency traders and find that some professional currency traders are adept at moderating losses.

We define drawdown as the maximum daily loss, proxied by the daily percentage return, for an individual currency trader. If top performing traders in this sample are skilled it is expected to mitigate their losses and thus have a lower drawdown than worst performing traders. Table 10 presents the drawdown results. Panel A reports the full sample results for all 428 account and quartile rankings based on the significance of alpha from the four-factor currency model. Panel B presents the results for account holders with an age over 80 days with similar rankings on performance. We also report the difference in means between the top performers in Q1 and the worst performers in Q4 for both the full sample and the age-truncated sample.

The results for the full sample of 428 accounts show that the full-sample mean drawdown is -16.81 percent. The quartile ranks show that the top performers (Q1) have a mean daily drawdown of -16.07 percent. This is lower than the worst performers (Q4), who have a mean daily drawdown of -19.15 percent and also lower than Q3 currency traders who have a mean daily drawdown of -16.84 percent. It is notable that Q2, which contains the second highest group of currency performers, has a lower drawdown of -15.19 percent than the top performers (-16.81 percent). Despite the differences between the groups, the differences are not statistically significant. The top performers in Q1 have a drawdown that is 3.08 percent lower than the worst performing traders in Q4 yet it is not statistically significant (t-statistic=1.29).

****Insert Table 10 about here****

In summary, the analysis of individual trades shows that a sizable percentage of traders in this sample are able to beat the odds and earn a profit on their trades, significantly different from pure chance. This implies these traders possess skill. Furthermore, the analysis of drawdowns reveals that the top performing traders have a better ability to mitigate downside losses than the worst performing traders, yet the difference is not significant.

5.8.3 Skill Measured by Timing Ability

Melvin and Shand (2011), argue that the ability of currency traders to time their exposure to systematic factors is an important contribution to performance. Melvin and Shand (2011) examine the returns of professional currency traders and find there is some evidence of timing ability amongst professional currency managers. Specifically, they show that out of the 42 currency managers they analyze, 13 timed the carry trade, 5 timed the PPP, and 9 timed momentum.

Similarly, if individual currency traders in this sample possess skill, they should also possess timing abilities. Our final inquiry of skill explores the ability of these individual currency traders to time the currency factors (*Carry*, *Mom* and *Value*). To test the timing abilities of the individual currency traders in this sample we follow a similar approach to Melvin and Shand (2011) and estimate the following equation:

$$r_{j,t} = \alpha_j + \sum_{i=1}^{3} \beta_{i,t}\left[F_{i,t} \middle| F_{i,t} > 0\right] + \sum_{i=1}^{3} \gamma_{i,t}\left[F_{i,t} \middle| F_{i,t} < 0\right] \tag{11}$$

where r is the return of individual currency trader j at time t; F is the return associated with factor i, and the factors are decomposed into positive and negative return observations. Individual currency trader timing ability is inferred from whether they load positive (negatively) on the factors when factor returns are positive (negative). We estimate this regression for all 428 accounts. For the sake of brevity, and since our main inquiry is whether the coefficients are significant, we report a summary of all significant coefficients (at the 5 percent level of significance) in Table 11. Full sample results are available upon request from the authors.

The results of the timing model, as shown in Table 11, support that some traders in this sample possess skill at timing the factors. Specifically, we find that 36 individual currency traders (8.41 percent) timed the carry trade. This implies that 36 traders have skill at timing the carry trade when the carry trade earned a positive daily return ("CarryPos"). 54 individual currency traders have significant coefficients when the carry factor earns negative returns ("CarryNeg"). This reveals a sizable percentage, 12.62 percent of all individual traders, have the ability to successfully time the carry trade when it earns negative returns.

To summarize the remainder of the results of the timing model presented in Table 11 we also report the total percentage of each factor coefficient that is significant in the final column. The coefficient with the lowest total percentage of significance is negative momentum ("MomNeg") at 7.71 which suggests that 33 out of 428 accounts were able to time momentum. CarryNeg has the highest percentage at 12.62 percent. It is notable that Melvin and Shand (2011) report that 5 out of 42 (approximately 11.9 percent) of traders successfully timed the PPP (referred to the value trade in this paper), 13 timed the carry (30.95 percent), and 9 timed momentum (21.4 percent). Here we report that approximately 17.29 percent of the individual traders successfully timed the value trade ("ValuePos" and "ValueNeg"), 17.52 percent timed momentum, and 21.03 percent timed carry. Although a direct comparison between the professional traders in Melvin and Shand (2011) and our sample warrants caution, our results demonstrate that some individual currency traders can time the factors implying that they have skill somewhat similar to professional currency traders.

****Insert Table 11 about here****

Table 2a. Full-Sample Value Weighted Results of the Daily Abnormal Return Measures for All Individual Currency Trader Accounts, 2004–2009

This table reports performance results for 428 individual currency traders at a proprietary online advisory service from March 2004 through September 2009. Performance measures are computed from daily gross and net returns, which are calculated from account records, and value-weighted portfolios are formed with the daily return data. Net returns account for a 3-pip ($3.00) transaction cost applied to each round trip transaction. Panel A presents results for the gross (net) return on equally weighted portfolios. Raw returns are calculated as the daily returns earned in aggregate by the account holders. Passive benchmark returns are calculated by subtracting the daily return of the DBCR from the daily raw return. The four-factor alpha is the intercept from the four-factor currency model of Pojarliev and Levich (2010b), where the excess equally weighted portfolio returns is regressed on four factors that mimic strategies used by professional currency traders: carry trade, momentum, PPP, and volatility. Excess returns are calculated by subtracting the daily LIBOR rates from the equally weighted portfolio return. Panel B sorts the account holders into performance quartiles. Ranks are calculated by four-factor alpha t-statistic rankings, with the top-performing accounts (with the highest alpha t-statistic) in quartile 1 and the lowest-performing currency traders in quartile 4. The t-statistics are in parentheses and significant values are bold; ** and * denote statistical significance at the 1% and 5% levels, respectively.

	Gross Returns			Net Returns		
	Raw Returns	Passive Benchmark	Four-Factor Alpha	Raw Returns	Passive Benchmark	Four-Factor Alpha
Panel A. Full-Sample Equal-Weighted Portfolio Performance Results						
	0.54	0.53	0.53	0.40	0.40	0.40
	(8.53)**	**(8.53)****	**(8.51)****	**(7.74)****	**(7.71)****	**(7.71)****
Panel B. Full-Sample Equal-Weighted Portfolio Results Sorted on Performance						
Q1 (Top performers)	2.67	2.66	2.67	1.25	1.24	1.24
	(18.03)**	**(17.98)****	**(18.03)****	**(10.90)****	**(10.86)****	**(10.82)****
Q2	0.33	0.32	0.31	0.75	0.75	0.76
	(2.44)**	**(2.45)***	**(2.40)***	**(6.20)****	**(6.20)****	**(6.25)****
Q3	-0.12	-0.12	-0.11	0.34	0.34	0.35
	(-1.49)	**(-1.45)**	**(-1.35)**	**(4.58)****	**(4.61)****	**(4.56)****
Q4 (Worst performers)	-1.27	-1.27	-1.28	-1.01	-1.01	-1.00
	(-17.77)**	**(-17.75)****	**(-17.81)****	**(-12.63)****	**(-12.63)****	**(-12.79)****
Panel C. Difference in Means Between Q1 and Q4						
Q1 – Q4	3.94	3.93	3.95	2.26	2.25	2.24
	(11.26)**	**(11.00)****	**(11.43)****	**(9.55)****	**(9.64)****	**(9.64)****

Table 8. Robustness Checks with Secondary Data Set of 74 Accounts from July 2010 to August 2011

This table reports summary statistics, performance results and trade activity results for 74 individual currency traders at a proprietary online advisory service from July 2010 to August 2011. Panel A reports mean daily returns, trades per day and the age of accounts. Trades per day for each account are calculated by dividing the total number of trades executed by account *i* over its account life, divided by the life of account *i* measured in days. The age of the account is measured in days. Panel B reports tertile sorts on gross performance and the difference in means between the top performers (T1) and the worst performers (T3). Panel C reports the results of sorts on trade activity, proxied my mean trades per day. Account holders are sorted into tertiles based on the mean number of trades executed for each trading day. Tertile 1 contains the account holders with the highest mean number of trades executed per day, and tertile 3 contains those with the lowest mean number of trades executed per day. The difference in means between the most active traders (T1) and the least active traders (T3) are also reported. The t-statistics are in parentheses and significant values are bold; ** denotes statistical significance at the 1% level.

A. Descriptive Statistics of Returns, Trade Activity and Age of Accounts

	Mean	25th Percentile	Median	75th Percentile	Obs.
Daily Gross Return	0.357	0.005	0.138	0.648	74
Trades Per day	2.35	0.64	1.69	3.14	74
Age (days)	201.30	133.00	171.50	266.00	74

B. Full Sample Results of Gross Returns with Sorts on Trade Activity

	Mean Gross Return	25th Percentile	Median	75th Percentile	Obs.
T1 (Best Performers)	1.154	0.648	1.073	1.302	25
T2	0.174	0.093	0.119	0.265	25
T3 (Worst Performers)	-0.283	-0.276	-0.125	0.003	24
Diff. Q1 - Q3	1.44				
	(8.37)**				

C. Full Sample Results of Gross Returns with Sorts on Trade Activity

Item	Mean Trades Per Day	Mean Gross Return	Obs.
T1 (Most Active Traders)	4.873	0.628	25
T2	1.544	0.367	25
T3 (Least Active Traders)	0.554	0.063	24
Diff. Q1 - Q3	4.32	0.57	
	(7.26)**	**(3.45)****	

Table 9. Skill based on the percentage of winning trades

This table reports the percentage of winning trades for all 428 accounts from 2004 to 2009. The percentage of winning trades is calculated as the total number of winning trades, defined as a trade with a net profit greater than zero, divided by the total number of trades for each account. Profit/Loss per trade is the USD gain or loss per trade executed by an individual currency trader. Panel A reports the results for the full sample of 428 accounts. Panel B reports the percentage of winning trades and profit/loss based on performance sorts, where quartile 1 contains the top-performing currency traders and quartile 4 contains the worst-performing traders. Each quartile contains 107 accounts. Panel C reports the difference in means between quartiles 1 and 4. The t-statistics are reported in parentheses and test whether the percentage of winning trades is significantly different from 50 percent. Significant values are bold, and * denotes statistical significance at the 1 percent level.

Panel A - Percentage of Winning Trades and Mean Profit/Loss Per Trade for Full Sample

	Percentage of Winning Trades				Profit/Loss Per Trade (USD)	
	Mean	Std Dev	Minimum	Maximum	Obs.	Mean
Full Sample	53.97	20.13	4.05	100	79042	29.85
	(4.08)*					**(2.56)***

Panel B - Percentage of Winning Trades and Mean Profit/Loss Per Trade Sorted on Performance

	Mean	Std Dev	Minimum	Maximum	Obs	Mean
1(Top performers)	66.78	17.99	21.36	100	10744	248.31
	(9.65)*					**(16.43)***
2	58.50	18.95	8.89	100	14947	69.58
	(4.64)*					**(2.71)***
3	48.33	16.61	11.76	88.64	19199	-66.71
	(1.04)					**(-2.26)***
4(Worst performers)	42.26	17.81	4.05	79.71	34152	-601.21
	(4.50)*					**(-17.42)***

Panel C Difference in Means of Winning Trades

	Mean Diff.					Mean Diff.
Q1 less Q 4	24.53					849.51
	(4.50)*					**(25.41)***

Table 10. Drawdown Proxied by Largest One-Day Percent Decline

This table reports drawdown for all 428 accounts from 2004 to 2009. Drawdown is calculated the largest daily negative return for an accountholder. Panel A reports the results for the full sample of 428 accounts and for quartile ranks based on the statistical significance of alpha from Pojarliev and Levich (2010b) four-factor currency model where quartile 1 contains the top-performing currency traders and quartile 4 contains the worst-performing traders. t-statistics are reported in parentheses.

Panel A. Full Sample Results

	Largest Daily Percentage Decline		
	Mean	Std. Dev.	Obs.
Full Sample	-16.81	16.45	428
Q1 (Top performers)	-16.07	16.45	107
Q2	-15.19	15.50	107
Q3	-16.84	15.11	107
Q4 (Worst performers)	-19.15	18.48	107
Diff Q1-Q4	3.08		
	(1.29)		

Table 11. Summary of Statistically Significant Coefficients for the Timing Model

This table reports the results of the timing model for regressions of all 428 individual currency trading accounts. The timing model is defined as: $r_{j,t} = \alpha_j + \sum_{i=1}^{3} \beta_{i,t} [F_{i,t} | F_{i,t} > 0] + \sum_{i=1}^{3} \gamma_{i,t} [F_{i,t} | F_{i,t} < 0]$ where r is the return of individual currency trader j at time t; F is the return associated with factor i, and the factors are decomposed into positive and negative return observations. Individual currency trader timing ability is inferred from whether they load positive (negatively) on the factors when factor returns are positive (negative). CarPos (CarryNeg), ValuePos (ValueNeg), and MomPos (MomNeg) are the explanatory variables in the timing model when the daily return for the carry, value and momentum is positive (negative). These variables are then regressed on daily net returns of individual currency traders. The number of statistically significant coefficients, at the 5 percent level of significance, are reported below. Full sample results for all 428 accounts are available upon request from the authors.

Variable/Factors	Number of Significant Coefficients	percent
CarryPos	36	8.41%
CarryNeg	54	12.62%
ValuePos	38	8.88%
ValueNeg	36	8.41%
MomPos	42	9.81%
MomNeg	33	7.71%