Appendix to
"Short Interest and Aggregate Stock Returns"

February 10, 2016

This appendix provides complete results for the robustness checks discussed in the paper. Below, we briefly describe the contents of the appendix tables.

- Table A1: out-of-sample R^2 statistics for alternative detrending methods discussed in footnote 22 of the paper
- Table A2: IVX-Wald and $\widehat{\mathrm{qLL}}$ statistics discussed in Sections 3.4.1 and 3.4.2, respectively, of the paper
- Table A3: results for decadal subsamples discussed in Section 3.4.2 of the paper
- Table A4: results for different excess return measures discussed in Section 3.4.3 of the paper
- Table A5: results for different short interest constituents and quasi-value-weighted short interest discussed in Sections 3.4.4 and 3.4.5, respectively, of the paper

Table A1
Out-of-sample R^2 statistics for alternative detrending methods, 1990:01–2014:12.

The table reports the proportional reduction (in percent) in mean squared forecast error (MSFE) at the h-month horizon for a predictive regression forecast of the S&P 500 log excess return based on SII and the detrending method in the first column vis-á-vis the prevailing mean benchmark forecast, where statistical significance is based on the Clark and West (2007) statistic for testing the null hypothesis that the prevailing mean MSFE is less than or equal to the predictive regression MSFE against the alternative hypothesis that the prevailing mean MSFE is greater than the predictive regression MSFE; *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. SII is computed as the deviation in the log of EWSI from a linear, quadratic, or cubic trend (as indicated in the first column), where EWSI is the equal-weighted mean across all firms of the number of shares held short in a given firm normalized by each firm's shares outstanding. "Stochastic" indicates that SII is computed as the deviation in the log of EWSI from a 60-month backward-looking moving average.

(1)	(2)	(3)	(4)	(5)
Detrending method	$h=1$	$h=3$	$h=6$	$h=12$
Linear	1.94***	6.54***	11.70***	13.24**
Quadratic	1.45***	5.74***	10.36***	12.62***
Cubic	0.39	2.78**	6.08**	−0.76
Stochastic	0.96**	4.28***	8.16***	9.89***

Table A2

Kostakis, Magdalinos, and Stamatogiannis (2015) and Elliott and Müller (2006) test results, 1973:01–2014:12.

The table reports test results concerning β for the predictive regression model,

$$r_{t:t+h} = \alpha + \beta \text{SII}_t + \varepsilon_{t:t+h} \text{ for } t = 1, \ldots, T-h,$$

where $r_{t:t+h} = (1/h)(r_{t+1} + \cdots + r_{t+h})$, r_t is the S&P 500 log excess return for month t, and SII_t is the short interest index. SII is computed as the deviation in the log of EWSI from a linear trend, where EWSI is the equal-weighted mean across all firms of the number of shares held short in a given firm normalized by each firm's shares outstanding. IVX-Wald is the Kostakis, Magdalinos, and Stamatogiannis (2015) Wald statistic for testing H_0: $\beta = 0$ against H_A: $\beta \neq 0$; the 10%, 5%, and 1% critical values are 2.71, 3.84, and 6.64, respectively. $\widehat{\text{qLL}}$ is the Elliott and Müller (2006) statistic for testing H_0: $\beta_t = \beta$ for all t; the 10%, 5%, and 1% critical values are −7.14, −8.36, and −11.05, respectively (where we reject for small values). *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$h = 1$		$h = 3$		$h = 6$		$h = 12$	
IVX-Wald	$\widehat{\text{qLL}}$	IVX-Wald	$\widehat{\text{qLL}}$	IVX-Wald	$\widehat{\text{qLL}}$	IVX-Wald	$\widehat{\text{qLL}}$
3.38^{*}	−3.72	4.51^{**}	−4.86	4.60^{**}	−4.91	3.67^{*}	−5.02

Table A3

In-sample predictive regression estimation results for decadal subsamples.

The table reports the ordinary least squares estimate of β and R^2 statistic for the predictive regression model,

$$r_{t:t+h} = \alpha + \beta \text{SII}_t + \varepsilon_{t:t+h} \;\text{ for }\; t = 1, \dots, T-h,$$

and subsample given in the first column, where $r_{t:t+h} = (1/h)(r_{t+1} + \cdots + r_{t+h})$, r_t is the S&P 500 log excess return for month t, SII_t is the short interest index, and $(-)$ indicates that we take the negative of SII_t. SII is computed as the deviation in the log of EWSI from a linear trend estimated for 1973:01 to 2014:12. EWSI is the equal-weighted mean across all firms of the number of shares held short in a given firm normalized by each firm's shares outstanding. SII is standardized to have a standard deviations of one. Brackets below the $\hat{\beta}$ estimates report heteroskedasticity- and autocorrelation-robust t-statistics for testing H_0: $\beta = 0$ against H_A: $\beta > 0$; *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively, according to wild bootstrapped p-values.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	$h=1$		$h=3$		$h=6$		$h=12$	
Subsample	$\hat{\beta}$	R^2 (%)	$\hat{\beta}$	R^2 (%)	$\hat{\beta}$	R^2 (%)	$\hat{\beta}$	R^2 (%)
1973:01–1982:12 $(-)$	0.68	1.71	0.72	5.82	0.65	9.04	0.77	28.10
	[1.97]**		[2.56]**		[2.34]*		[3.53]**	
1983:01–1992:12 $(-)$	0.84	1.28	0.65	2.39	0.53	4.08	0.49	9.53
	[1.14]		[1.03]		[1.34]		[1.79]	
1993:01–2002:12 $(-)$	0.36	0.18	0.63	1.74	0.53	2.28	−0.37	1.56
	[0.49]		[0.92]		[0.73]		[−0.45]	
2003:01–2014:12 $(-)$	0.62	4.50	0.72	13.83	0.79	25.37	0.76	41.47
	[2.42]***		[2.68]***		[2.60]**		[3.10]**	

Table A4

In-sample predictive regression estimation results for different excess return measures, 1973:01–2014:12.

The table reports the ordinary least squares estimate of β and R^2 statistic for the predictive regression model,

$$r_{t:t+h} = \alpha + \beta \mathrm{SII}_t + \varepsilon_{t:t+h} \text{ for } t = 1, \dots, T-h,$$

where $r_{t:t+h} = (1/h)(r_{t+1} + \cdots + r_{t+h})$, r_t is one of four possible measures of the log excess return in the first column for month t, SII_t is the short interest index, and $(-)$ indicates that we take the negative of SII_t. SII is computed as the deviation in the log of EWSI from a linear trend. EWSI is the equal-weighted mean across all firms of the number of shares held short in a given firm normalized by each firm's shares outstanding. SII is standardized to have a standard deviation of one. Brackets below the $\hat{\beta}$ estimates report heteroskedasticity- and autocorrelation-robust t-statistics for testing H_0: $\beta = 0$ against H_A: $\beta > 0$; *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively, according to wild bootstrapped p-values.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	$h=1$		$h=3$		$h=6$		$h=12$	
Log excess return	$\hat{\beta}$	R^2 (%)	$\hat{\beta}$	R^2 (%)	$\hat{\beta}$	R^2 (%)	$\hat{\beta}$	R^2 (%)
Value-weighted S&P 500 $(-)$	0.50	1.24	0.56	4.37	0.57	8.07	0.53	12.89
	[2.50]***		[2.88]***		[2.73]**		[2.70]**	
Equal-weighted S&P 500 $(-)$	0.61	1.40	0.69	4.72	0.68	9.04	0.62	16.70
	[2.46]***		[2.92]***		[2.90]***		[3.27]***	
Value-weighted CRSP $(-)$	0.52	1.25	0.58	4.23	0.58	7.77	0.53	12.42
	[2.41]***		[2.72]***		[2.62]**		[2.61]**	
Equal-weighted CRSP $(-)$	0.69	1.45	0.80	4.52	0.78	8.34	0.68	13.27
	[2.49]***		[2.93]***		[3.04]***		[3.12]***	

Table A5

In-sample predictive regression estimation results for different short interest measures, 1973:01–2014:12.

The table reports the ordinary least squares estimate of β and R^2 statistic for the predictive regression model,

$$r_{t:t+h} = \alpha + \beta \mathrm{SII}_t + \varepsilon_{t:t+h} \ \text{ for } \ t = 1, \ldots, T-h,$$

where $r_{t:t+h} = (1/h)(r_{t+1} + \cdots + r_{t+h})$, r_t is the S&P 500 log excess return for month t, SII_t is the short interest index, and ($-$) indicates that we take the negative of SII_t. SII is computed as the deviation in the log of one of the four possible short interest measures in the first column from a linear trend. EWSI is the equal-weighted mean across all firms of the number of shares held short in a given firm normalized by each firm's shares outstanding. EWSI includes common equities, ADRs, ETFs, and REITs. QVWSI is the value-weighted mean across all firms of the number of shares held short in a given firm normalized by each firm's shares outstanding, where the weights are based on log market capitalization. QVWSI includes common equities, ADRs, ETFs, and REITs. SII is standardized to have a standard deviation of one. Brackets below the $\hat{\beta}$ estimates report heteroskedasticity- and autocorrelation-robust t-statistics for testing H_0: $\beta = 0$ against H_A: $\beta > 0$; *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively, according to wild bootstrapped p-values.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	$h=1$		$h=3$		$h=6$		$h=12$	
Short interest measure	$\hat{\beta}$	R^2 (%)	$\hat{\beta}$	R^2 (%)	$\hat{\beta}$	R^2 (%)	$\hat{\beta}$	R^2 (%)
EWSI, all assets ($-$)	0.50	1.24	0.56	4.37	0.57	8.07	0.53	12.89
	[2.50]***		[2.88]***		[2.73]**		[2.70]**	
EWSI, common equities only ($-$)	0.41	0.82	0.43	2.60	0.43	4.78	0.41	8.05
	[1.96]**		[2.11]**		[1.91]*		[1.82]*	
EWSI, excluding common equities ($-$)	0.41	0.82	0.42	2.49	0.42	4.52	0.40	7.61
	[1.93]**		[2.03]**		[1.83]*		[1.73]	
QVWSI, all assets ($-$)	0.49	1.17	0.55	4.15	0.55	7.72	0.52	12.38
	[2.43]***		[2.81]***		[2.66]**		[2.64]**	

References

Clark, T.E., West, K.D., 2007. Approximately normal tests for equal predictive accuracy in nested models. Journal of Econometrics 138, 291–311.

Elliott, G., Müller, U.K., 2006. Efficient tests for general persistent time variation in regression coefficients. Review of Economic Studies 73, 907–940.

Kostakis, A., Magdalinos, T., Stamatogiannis, M.P., 2015. Robust econometric inference for stock return predictability. Review of Financial Studies 28, 1506–1553.