Economics Bulletin

Volume 43, Issue 2

Testing for herding using different return definitions: a comparison between simple and logarithmic returns

Junkai Wang
University of Hull

Robert Hudson
University of Hull

Abstract

We outline theoretically why different return measures (simple or logarithmic) may affect tests for herding. Comprehensive empirical tests using data from many major world financial markets confirm that different returns measures do frequently lead to different conclusions about the presence of herding.

Citation: Junkai Wang and Robert Hudson, (2023) "Testing for herding using different return definitions: a comparison between simple and logarithmic returns", *Economics Bulletin*, Volume 43, Issue 2, pages 1070-1080
Contact: Junkai Wang - robert.hudson@hull.ac.uk, Robert Hudson - robert.hudson@hull.ac.uk.
Submitted: October 19, 2022. **Published:** June 30, 2023.

1. Introduction

Herding in financial markets occurs when investors are guided by the collective behaviour of other market participants rather than their own beliefs. Numerous studies have empirically investigated herding in many world markets[1]. Empirical investigations of herding necessarily need to use measures of the returns on the assets involved. Two common ways to measure returns are used in financial theory and practice: logarithmic and simple returns[2]. Herding investigations provide an interesting setting to investigate the possible importance of using different asset return measures. The use of different return measures can sometimes lead to significantly different outcomes (Hudson and Gregoriou, 2015). To our knowledge the effect of using different return measures has not previously been investigated in the context of herding and indeed many papers in the literature do not indicate which measure is used. The return measure used may be particularly important in herding tests for two reasons. Firstly, herding is often proposed to be particularly prevalent when there are large price movements in markets which is when the logarithmic and simple measures have their greatest differences. Secondly, the most common test for herding tests for non-linear relationships between market returns and the dispersion of the returns on the assets in the market. Given there is a non-linear relationship between logarithmic and simple returns, the use of the different return measures is likely to lead to different conclusions from this test.

In this paper we initially outline why the return measure used may affect tests for herding. We then carry out comprehensive empirical tests using different data sets from many of the world's major financial markets to show that different returns measures can frequently lead to different conclusions about the presence of herding.

2. Effect of the Return Measure on Tests for Herding

Most empirical work on herding uses a test developed by Chang, Cheng and Kohrana (2000) which we subsequently refer to as the CCK test. Due to its popularity, we focus on this test in our investigation. The CCK test is based on the proposition that the cross-sectional absolute deviation of stock returns (CSAD) should be linearly related to overall market returns where:

$$CSAD_t = \frac{1}{N}\sum_{i=1}^{N} \left| R_{¿} - R_{mt} \right| \qquad (1)$$

Given this, the CCK approach for testing for herding is to examine a regression model of the following basic form:

---

[1] See, for example, Christie and Huang (1995) and Clements, Hurn and Shi (2017) for the US market, Economou, Kostakis and Philippas (2011) and Galariotis, Krokida and Spyrou (2016) for European financial markets and Demirer and Kutan (2006) and Arjoon, Bhatnagar and Ramlakhan (2020) for Asian markets.

[2] The log return is defined by $R_{¿} = \ln\left(\frac{P_t}{P_{t-1}}\right)$. The simple return is defined by $R_{St} = ¿(\frac{P_t}{P_{t-1}} - 1 ¿$ where P_t is the stock price at time t.

$$CSAD_t = \alpha + \gamma_1 R_{m,t} + \gamma_2 \left| R_{m,t} \right| + \gamma_3 R_{m,t}^2 + \varepsilon_t \qquad (2)$$

If γ_3 is negative and significant that indicates herding as $CSAD_t$ increases less than proportionately to $R_{m,t}$ which is in accordance with market participant imitating one another as market movements become larger. Conversely, if γ_3 is positive and significant that indicates anti-herding as $CSAD_t$ increases more than proportionately to $R_{m,t}$ as market participant have less tendency to imitate one another as market movements become larger. If γ_3 is not statistically significant that is taken as evidence of no herding. Theoretically, it is often proposed that herding is more likely in times of market stress so often Equation 2 is used to examine periods of large absolute market movements either by modifying the equation with appropriate dummy variables or by looking at subsets of the data. For example, Chiang, Li and Tan (2010) used quantile regression to estimate herding behaviour in Chinese stock markets, showing that there tends to be more significant herding in the quantiles associated with larger returns. We adopt the approach of looking at appropriate subsets of the data in this paper.

If a study uses log or simple returns throughout there is no theoretical issue preventing $CSAD_t$ from being proportionate to $R_{m,t}$. However, there is not a linear relationship between log and simple returns for a set of price movements. In particular, if we consider the Taylor series expansion of $R_{¿}$:

$$R_{¿} = \ln\left(1 + R_{St}\right) = R_{St} - \frac{{R_{St}}^2}{2} + \frac{{R_{St}}^3}{3} - \frac{{R_{St}}^4}{4} + \ldots$$

Thus, if there is a precisely proportionate mathematical relation between $CSAD_t$ and $R_{m,t}$ for simple (log) returns that cannot also be true for log (simple) returns. Thus, findings regarding the presence of herding will, necessarily, depend on the return measure used. We assess the empirical significance of any differences in the section below.

Table 1 numerically demonstrates this effect and some of its implications. For simple returns, $CSAD_S$ is a constant precise linear function of $\left|R_{Sm}\right|$ so in this case, by definition, there is no herding[3]. Without loss of generality, we assume two equally weighted stocks in the market. $R_{S1t} = 1.5\ R_{Smt}$ and $R_{S2t} = 0.5\ R_{Smt}$.

[3] To show that $CSAD_S$ is a constant precise linear function of $\left|R_{Sm}\right|$ consider a market consisting of two equally weighted stocks S1 and S2. Without loss of generality, we assume $R_{S1t} = 1.5\ R_{Smt}$ and $R_{S2t} = 0.5\ R_{Smt}$ Then

$R_{sm} = 0.5\, R_{s1} + 0.5\, R_{s2}$

From Eqn 1: $CSAD_s = 0.5\left(\left|R_{s1} - R_{sm}\right| + \left|R_{s2} - R_{sm}\right|\right)$

We have $R_{s2} = 2\, R_{sm} - R_{s1}$

Then:

$$CSAD_s = 0.5\left(\left|R_{s1} - R_{sm}\right| + \left|R_{sm} - R_{s1}\right|\right)$$
$$¿ \left|R_{s1} - R_{sm}\right|$$

In our example, $R_{S1} = 1.5\ R_{Sm}$

So that $CSAD_s = ¿\, 0.5\, R_{sm} \vee ¿ = 0.5 \vee R_{sm} \vee ¿$ i.e. $CSAD_s$ is a constant precise linear function of $\left|R_{Sm}\right|$.

In contrast, for log returns, $CSAD_L$ is not a linear function of $\left|R_{Lm}\right|$. We can observe that for negative market returns $CSAD_L$ has a convex relationship with $\left|R_{Lm}\right|$ which is associated with anti-herding, conversely for positive market returns $CSAD_L$ has a concave relationship with $\left|R_{Lm}\right|$ which is associated with herding. Thus, using different return measures may give different results regarding the presence of herding and these results are also somewhat dependent on market conditions. The relationship between tests for herding using the different measures will depend on whether the overall market is rising or falling and on the magnitude of market movements.

The example in Table 1 is deterministic so does not allow for the stochastic deviations from the underlying model that are observed in real stock markets and the returns have been chosen for illustrative purposes. This gives rise to the important practical issue of whether the use of different return measures can produce substantially different conclusions using actual market data. We deal with this question in section 3.

Table 1: Numerical demonstration showing how $CSAD_S$ and $CSAD_L$, the cross-sectional absolute deviation of simple and logarithmic stock returns respectively, change in response to changing market returns for a market of two assets. All figures reported are as at time t. R_s and R_L denote the simple and logarithmic returns on assets.						
Asset	R_S	R_L	$CSAD_S$	$CSAD_L$	$\frac{CSAD_S}{\lvert R_S \rvert}$	$\frac{CSAD_L}{\lvert R_L \rvert}$
1	-0.75	-1.38629				
2	-0.25	-0.28768				
Marke t	-0.5	-0.83699	0.25	0.549306	0.5	0.656289
1	-0.6	-0.91629				
2	-0.2	-0.22314				
Marke t	-0.4	-0.56972	0.2	0.346574	0.5	0.608326
1	-0.45	-0.59784				
2	-0.15	-0.16252				
Marke t	-0.3	-0.38018	0.15	0.217659	0.5	0.572519
1	-0.3	-0.35667				
2	-0.1	-0.10536				
Marke t	-0.2	-0.23102	0.1	0.125657	0.5	0.543929
1	-0.15	-0.16252				
2	-0.05	-0.05129				
Marke t	-0.1	-0.10691	0.05	0.055613	0.5	0.520202
1	0.15	0.139762				
2	0.05	0.04879				
Marke t	0.1	0.094276	0.05	0.045486	0.5	0.482476
1	0.3	0.262364				
2	0.1	0.09531				
Marke t	0.2	0.178837	0.1	0.083527	0.5	0.467056
1	0.45	0.371564				
2	0.15	0.139762				
Marke t	0.3	0.255663	0.15	0.115901	0.5	0.453335
1	0.6	0.470004				

	2	0.2	0.182322				
Market		0.4	0.326163	0.2	0.143841	0.5	0.44101
	1	0.75	0.559616				
	2	0.25	0.223144				
Market		0.5	0.39138	0.25	0.168236	0.5	0.429854

3. Empirical Evidence from Global Markets of the Effect of Using different Return Measures.

In this section we present the results of herding tests for a number of major world markets based on the traditional CCK using both Logarithmic and Simple returns to see the difference between the measures. Robust regression methods are used throughout.

The data set is constructed from companies in the leading indices of Denmark (OMXC-20), Finland (HEX-25), US Dow Jones Composite, Germany (DAX-30), France (CAC-40), Greece (ATHEX), Italy (FTSE-MIB), Norway (OBX), Portugal (PSI-20), Spain (IBEX-35), Sweden (OMXS-30), Hong Kong Heng SENG as well as the UK market (FTSE-100). The data sample period is collected from Bloomberg over the period from 02/Jan/2002 to 31/May/2018.

Table 2 reports the descriptive statistics for the equally weighted average market return and the CCK measurements for each of the total thirteen different countries using logarithmic and simple returns. We see that, for each country, the difference between the two types of return are quite modest for standard deviation and CSAD whereas they are quite substantial for mean returns and the maximum and minimum returns.

Table 3 reports the results when Equation 2 is applied to the full data and subsets of the data based on absolute return size which are shown in Panels A and B respectively. As discussed above, the key determinant of herding or anti-herding is the sign and significance of the coefficient of $R^2_{m,t}$ so to conserve space this is the coefficient we report. For the countries in the full data set, for both log and simple returns, this coefficient is always positive which is associated with anti-herding. As the size of the data set is reduced by focusing on larger absolute returns, in accordance with theoretical expectations, the coefficient is much more likely to be negative which is associated with herding.

The key purpose of our paper is to compare the results for logarithmic and simple returns. For logarithmic returns on the full data, shown in Panel A, set we see that the coefficient of $R^2_{m,t}$is significant for 8 countries. For simple returns we see that the statistical significance of the coefficient changes for three countries. Denmark and Norway both become significant when they were not significant previously and the US loses its significance.

As we investigate the data sets that focus on larger absolute returns, shown in Panel B, we continue to see substantial numbers of countries with the relevant coefficient changing significance as we move from log to simple returns. For the regressions based on the sets of returns within the top 50% of absolute returns or more we see that the coefficients for at least 3 and in most cases 4 of the countries change significance. It is interesting that the changes are not in a systematic direction so that changing from logarithmic to simple returns may either increase or decrease to likelihood of finding evidence of herding.

Thus, the results regarding the presence of herding alter with a change in return calculation for around one quarter of the countries investigated for each of the data sets. We can consider the statistical significance of these numbers. Given conservative methodology we see that 3 out of 13 countries changing significance is of borderline statistical significance and 4 out of 13 countries changing significance is statistically significant[4].

[4] We can conceptualize the situation as having done tests for herding/anti-herding on 13 counties using logarithmic returns with the conclusion for each country being correct with 90% probability (assuming 10% significance level). Given this, if we repeat the tests using simple returns we can assess if the results are statistically significant by assuming the number of changes follow a binomial distribution with a 10% probably of each country changing significance. The probability of three or more changes is 13.39% and the probability of four or more changes is 3.42%. These conclusions are conservative as many of the initial logarithmic results are significant at over the 10% level.

Table 2: Descriptive data

Country	Variable	Mean	sd	Min	Max	Mean CSAD	N
Denmark	Log $R_{m,t}$	.044824	1.21182	-10.5563	7.99761	1.2098	4105
	Simple $R_{m,t}$	.069696	1.20643	-9.93637	8.38672	1.21065	
US	Log $R_{m,t}$	.030009	1.20649	-8.06138	9.54237	.908568	4132
	Simple $R_{m,t}$	.047245	1.20915	-7.69683	10.0664	.908986	
Finland	Log $R_{m,t}$	.029096	1.45997	-8.92102	8.93025	1.16771	4124
	Simple $R_{m,t}$	.053756	1.46356	-8.46355	9.37088	1.17074	
France	Log $R_{m,t}$	.020807	1.45872	-9.31602	8.91817	1.0055	4202
	Simple $R_{m,t}$	.042586	1.45982	-8.84619	9.38817	1.00639	
Germany	Log $R_{m,t}$	.022635	1.41664	-9.02234	11.1545	1.03381	4171
	Simple $R_{m,t}$	.045037	1.41719	-8.52537	11.8836	1.03473	
Greece	Log $R_{m,t}$	-.01962	1.66795	-15.9129	12.6811	1.82591	4063
	Simple $R_{m,t}$	.032309	1.66244	-14.0175	13.8705	1.82756	
HK	Log $R_{m,t}$	.041765	1.40087	-12.413	11.4602	1.15378	4050
	Simple $R_{m,t}$	.066823	1.40287	-11.5609	12.2546	1.15651	
Italy	Log $R_{m,t}$	.004424	1.41339	-8.56588	9.27357	1.10248	4168
	Simple $R_{m,t}$	.028636	1.41354	-8.14261	9.82029	1.10372	
Norway	Log $R_{m,t}$	.026429	1.83862	-12.3905	10.4173	1.50196	4120
	Simple $R_{m,t}$	.064272	1.86264	-11.9357	11.1138	1.5208	
Portugal	Log $R_{m,t}$	.007854	1.1991	-7.98493	8.74228	1.16989	4194
	Simple $R_{m,t}$	.029709	1.19824	-7.55258	9.39527	1.17161	
Spain	Log $R_{m,t}$	.017445	1.31686	-8.06577	9.71678	.977813	4174
	Simple $R_{m,t}$	.036884	1.31646	-7.69075	10.3766	.97829	
Sweden	Log $R_{m,t}$	.030535	1.61524	-9.30306	13.0496	1.01083	4123
	Simple $R_{m,t}$	.055766	1.61765	-8.82834	14.0028	1.01093	
UK	Log $R_{m,t}$	.020814	1.18132	-9.38468	7.88027	1.09882	4131
	Simple $R_{m,t}$	.04287	1.1798	-8.79727	8.34741	1.09861	

Table 3 – Results of herding tests – to conserve space only the coefficients of the $R^2_{m,t}$ term have been reported
Panel A – Full set of returns

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
	Denmark	US	Finland	France	Germany	Greece	Hong Kong	Italy	Norway	Portugal	Spain	Sweden	UK
$CSAD_t = \alpha + \gamma_1 R_{m,t} + \gamma_2 \lvert R_{m,t} \rvert + \gamma_3 R^2_{m,t} + \varepsilon_t$													
γ_3 Log	0.0328	0.0116*	0.0111**	0.0179***	0.0204***	0.0160***	0.0101	0.0223***	0.00122	0.00684	0.0175***	0.00721	0.0270***
	(1.62)	(1.76)	(2.00)	(4.74)	(2.60)	(5.35)	(1.49)	(2.89)	(0.23)	(1.47)	(4.22)	(1.59)	(3.26)
γ_3 Sim	0.0223**†	0.0105†	0.0140**	0.0185***	0.0195**	0.0156***	0.0109	0.0316**	0.0351**†	0.00546	0.0171***	0.00613	0.0284***
	(2.03)	(1.56)	(2.11)	(4.86)	(2.26)	(3.93)	(1.56)	(2.43)	(2.25)	(1.32)	(4.84)	(1.33)	(3.87)
3 Countries change significance													

*10% significance, **5% significance, 1% significance, t-stats shown in parentheses.
† Indicates whether the significance of the variable is different when it is calculated used simple rather than logarithmic returns.

Table 3 – Results of herding tests – to conserve space only the coefficients of the $R^2_{m,t}$ term have been reported
Panel B – Subsets of returns

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
	Denmark	US	Finland	France	Germany	Greece	Hong Kong	Italy	Norway	Portugal	Spain	Sweden	UK
Largest 50% of returns (50% of absolute value (above 25% and 25% below 0)) $CSAD_t = \alpha + \gamma_1 R_{m,t} + \gamma_2 \lvert R_{m,t} \rvert + \gamma_3 R^2_{m,t} + \varepsilon_t$													
γ_3 Log	0.0199	-0.00625	0.00745	0.00727	0.0116	0.00806*	-0.00166	0.0114	-0.0133*	0.00522	0.0104**	-0.00551	0.0151
	(0.75)	(-0.83)	(0.93)	(1.51)	(1.13)	(1.91)	(-0.27)	(1.07)	(-1.94)	(0.79)	(2.03)	(-1.28)	(1.58)
γ_3 Sim	0.00178	-0.00555	0.00820	0.00813*†	0.00961	0.00599†	0.000556	0.0225	0.0332†	0.00225	0.0105**	-0.00500	0.0160*†
	(0.15)	(-0.72)	(0.91)	(1.68)	(0.90)	(1.32)	(0.08)	(1.31)	(1.49)	(0.38)	(2.24)	(-1.20)	(1.88)
4 Countries change significance													
Largest 10% (10% of absolute value (above 5% and 5% below 0)) $CSAD_t = \alpha + \gamma_1 R_{m,t} + \gamma_2 \lvert R_{m,t} \rvert + \gamma_3 R^2_{m,t} + \varepsilon_t$													
γ_3 Log	0.00816	-0.0252*	0.0116	-0.0179*	-0.0154	0.00525	-0.0130*	-0.00025	-0.0267*	0.0305**	0.00384	-0.023***	-0.0156
	(0.19)	(-1.88)	(0.64)	(-1.74)	(-0.89)	(0.62)	(-1.72)	(-0.01)	(-1.65)	(2.24)	(0.32)	(-3.33)	(-1.07)
γ_3 Sim	-0.0298*†	-0.0227*	-0.00025	-0.0134†	-0.0163	0.000976	-0.0129†	0.0190	0.0502†	0.0254**	0.00974	-0.019***	-0.0132
	(-1.96)	(-1.72)	(-0.01)	(-1.28)	(-1.01)	(0.11)	(-1.50)	(0.63)	(1.05)	(2.00)	(1.04)	(-3.27)	(-0.89)
4 Countries change significance													
Largest 5% (5% of absolute value (above 2.5% and 2.5% below 0)) $CSAD_t = \alpha + \gamma_1 R_{m,t} + \gamma_2 \lvert R_{m,t} \rvert + \gamma_3 R^2_{m,t} + \varepsilon_t$													
γ_3 Log	-0.0189	-0.045***	-0.0183	-0.0294*	-0.0417**	0.0114	-0.031***	0.0186	-0.0257	0.0419**	-0.00158	-0.034***	-0.0454**
	(-0.39)	(-2.62)	(-0.69)	(-1.74)	(-2.12)	(1.04)	(-3.62)	(0.44)	(-1.01)	(2.12)	(-0.08)	(-4.18)	(-2.15)
γ_3 Sim	-0.074***†	-0.0370**	-0.0436	-0.0244†	-0.0420**	0.00179	-0.026***	0.0114	0.0540	0.0277†	0.0134	-0.030***	-0.0566**
	(-3.40)	(-2.20)	(-1.29)	(-1.35)	(-2.49)	(0.14)	(-2.61)	(0.28)	(0.72)	(1.56)	(1.00)	(-4.43)	(-2.42)
3 Countries change significance													
Largest 2% (2% of absolute value) $CSAD_t = \alpha + \gamma_1 R_{m,t} + \gamma_2 \lvert R_{m,t} \rvert + \gamma_3 R^2_{m,t} + \varepsilon_t$													
γ_3 Log	-0.0237	-0.0492*	-0.0751	-0.0496*	-0.0492	0.0215	-0.063***	0.0197	-0.0816**	-0.00442	-0.0196	-0.048***	-0.0968**
	(-0.44)	(-1.80)	(-1.23)	(-1.70)	(-1.47)	(1.26)	(-3.85)	(0.28)	(-2.12)	(-0.12)	(-0.58)	(-4.01)	(-2.33)
γ_3 Sim	-0.111**†	-0.0416	-0.107	-0.0209†	-0.0403	0.0202	-0.065***	-0.125	-0.0600†	-0.0209	-0.00842	-0.041***	-0.0992**
	(-2.37)	(-1.65)	(-1.60)	(-0.69)	(-1.49)	(0.86)	(-4.57)	(-1.24)	(-0.37)	(-0.85)	(-0.34)	(-4.19)	(-2.08)

4 Countries change significance
*10% significance, **5% significance, 1% significance, t-stats shown in parentheses. † Indicates whether the significance of the variable is different when it is calculated used simple rather than logarithmic returns.

4. Conclusions

In this paper we show both theoretically and empirically that different returns measures can lead to different conclusions about the presence of herding. In tests on many of the world's major financial markets we show that a change in research measure may lead to a different conclusion about the presence of herding in the region of a quarter of cases. Thus, in this area of finance research, we suggest that the choice of return measure to use in investigations is an important decision and should be carefully considered given the purpose of the study. Additionally, the return measure used should be reported as a matter of course to facilitate comparisons with other similar studies.

References

Arjoon, V., Bhatnagar, C. and Ramlakhan, P., (2020) "Herding in the Singapore Stock Exchange" *Journal of Economics and Business* **109**, 105889.

Chang, E. C., J. W. Cheng, and A. Khorana. (2000) "An Examination of Herd Behavior in Equity Markets: An International Perspective" *Journal of Banking and Finance* **24**1651–1679.

Chiang, T., Li, J. and Tan, L. (2010) "Empirical investigation of herding behavior in Chinese stock markets: Evidence from quantile regression analysis" *Global Finance Journal* **21(1)** 111-124.

Christie, W.G., Huang, R.D. (1995) "Following the pied piper: do individual returns herd around the market?" *Financial Analysts Journal* **51** 31-37.

Clements, A., Hurn, S. and Shi, S. (2017) "An empirical investigation of herding in the U.S. stock market" *Economic Modelling* **67** 184-192.

Demirer, R. and Kutan, A. (2006) "Does herding behavior exist in Chinese stock markets?" *Journal of International Financial Markets, Institutions and Money* **16(2)** 123-142.

Economou, F., Kostakis, A. and Philippas, N. (2011) "Cross-country effects in herding behaviour: Evidence from four south European markets" *Journal of International Financial Markets, Institutions & Money* **21(3)** 443–460.

Galariotis, E., Krokida, S. and Spyrou, S. (2016) "Herd behaviour and equity market liquidity: Evidence from major markets" *International Review of Financial Analysis* **48** 140-149.

Hudson, R.S. and Gregoriou, A. (2015) "Calculating and comparing security returns is harder than you think: A comparison between logarithmic and simple returns" *International Review of Financial Analysis* **38** 151-162.