Study of Mutual Fund Misreports

Dhiren Khodidas Charania

A dissertation presented to the faculty of the College of Science
in partial fullfilment of the requirement for the degree of
Master of Financial Engineering

UNIVERSITY OF CANTERBURY
Te Whare Wānanga o Waitaha

College of Science
University of Canterbury
New Zealand

Project Supervisors:

Professor Jedrzej Bialkowski

Associate Professor Carl Scarrott

May 2019

Acknowledgement

The success and final outcome of this project required a lot of guidance and assistance from my supervisor and I am extremely privileged to have got this all along the completion of my project. All that I have done is only due to such supervision and assistance and I would not forget to thank them.

I respect and thank Mr. Jedrzej Bialkowski, for providing me an opportunity to do the project work and giving me all support and guidance which made me complete the project duly. His guidance in the research with innumerable ideas was remarkable. I am thankful to him for providing such a nice support and guidance.

I owe my deep gratitude to Mr. Carl Scarrott, who took keen interest on my project work and guided me all along, till the completion of project work by providing all the necessary information for developing a good program and of course mathematics. He is extraordinarily generous with his time and energy and countless help and guidance.

Many thanks to PhD. student Cheng Zhan (Hannah) for her help by sharing past papers on Fama-French model and conversation.

Contents

1 Introduction

With the development of financial markets, the American financial system is dominated by institutional investors as the main players. Some of them enjoy low regulation when compared to other investment vehicles. Due to limited regulations on hedge funds, they played a crucial role in the financial crisis of 2007. In 2011, the SEC re-enacted a registration law to examine the impact of systematic risk of the financial system and to prevent hedge fund fraud (Dodd-Frank Wall Street Reform Act 2010). Researchers found anomalies in hedge funds returns (Bollen & Pool (2009), Agarwal, Daniel & Naik (2011)).

The anomaly in the form of discontinuity was found in monthly returns of hedge funds by Bollen & Pool (2009). They concluded that some mangers distort returns in liquid securities. They distort returns when funds returns are at managers' discretion and when their reported returns are not closely monitored. Bollen & Pool (2009) stated that hedge fund monthly returns are inclined to report marginally positive return as opposed to negative returns in order to make their overall returns look more attractive to investors and to prevent fund outflow.

Hedge funds are compensated by incentive fees based on annual performance exceeding prespecified thresholds. Agarwal, Daniel & Naik (2011) concluded that hedge funds inflate their reported returns in order to earn higher fees. They mentioned that December returns of hedge funds are significantly higher than their average returns from January to November. They also found that December spike is systematically related to the benefits and costs associated with return management. Hedge funds borrow from their January returns of the subsequent year to improve their December year. This can be achieved by funds pushing up the security prices at December-end by last minute buying, which is followed by price reversal in January.

Jorge Perez (2014) reported that some fund managers are involved in window dressing their portfolios to hide their actual investment strategy from their investors or from competing funds. Window dressing is disclosing disproportionately higher holdings in stocks that have done well over a reporting period. Thus, studies analysing only the returns of the disclosed holdings might be subject to significant measurement error, as they do not capture interim trades and various hidden costs. Extensive literature has shown hedge fund engage in return smoothing and misreporting returns to improve the overall performance of the fund for various reasons, including attracting potential investors and preventing investor outflow. SEC has been recently concerned about issues related to accurate security valuation in hedge funds.

Mutual funds are no exception. The effects in the form of anomalies in returns, involves managers altering or distorting their portfolio in an attempt to mislead investors by window dressing. If investors are misled by funds' window-dressing the funds are rewarded with higher flows. Conversely if investors are not deceived by window-dressing the funds, are punished with lower flows. The fund managers take advantage of flexibility in disclosing portfolio holdings with a delay of up to 60 days following quarter end. Agarwal, Gay & Ling (2014) concluded that managers who are less skilled and perform poorly are associated with window dressing. These poorly performing managers may decide to window dress towards quarter end and thus rebalance to disclose disproportionately higher and lower proportions of winner and loser stocks during quarter, respectively.

Hyunk-Suk Choi (2015) revealed that January is the month when equity funds experience the largest net cash flows and December is the month with the smallest net cash flows. The large net cash flows in January are attributed to increased purchases, and the small net flows in December are due to increased redemptions. The investors make asset allocation decisions more actively around the turn-of-the-year.

Brown, Harlow, & Starks (1996), Chevalier & Ellison (1997); Sirri & Tufano(1998); and Huang, Wei, & Yan (2007) suggests that many investors evaluate funds on a calendar-year basis, which may provide greater incentives to window dress in December. Haugen & Lakonishok (1988) and Ritter & Chopra (1989) argue that window dressing can potentially explain the January effect. It is important to note that only some funds appear to engage in window-dressing behaviour.

Window dressing is associated with trading which results in higher trade costs, such as trading costs, agency costs, and negative investor externalities. Such funds exhibits lower performance in subsequent quarters. Karperczyk, Sialm & Zheng (2008) have measured the impact of unobserved actions by comparing the actual mutual fund performance with the performance of a hypothetical portfolio that invests in previously disclosed fund holdings. The return gap is based on the comparison of the net investor return and the net return of the funds' holding. The impact of unobserved actions is included in the investor return but not in the returns of the hypothetical portfolio.

The return gap is negatively related to the hidden costs and positively related to the hidden benefits of a mutual fund. Funds with high past return gaps tend to perform consistently better before and after adjusting for differences in their risks and styles. The return gap is related to other fund attributes, such as size, age, and average new money growth (NMG). In short, funds that are subject to higher commissions will have higher hidden costs. Grinblatt & Titman (1989) use the difference between investor and holdings returns to estimate the total transactions costs for mutual funds. Interim trades within a quarter and possible window dressing activities may affect the estimated difference.

Before 2003, late hour trading was dominant in the mutual funds industry. Late hour trading allows selected customers to buy and sell after 4pm closing time at the price of 4 pm. This practise is performed by those customers who trade on the basis of information after 4pm to guide their trade. Such buying and selling also allows mutual funds to calculate net asset value (NAV) after an order is received. Late hour trading is illegal. Zitzewitz (2004) reported that the extent of late hour trading has dropped sharply since September 2003 when the main allegation of mutual fund scandal in 2003 was late hour trading. In my case, the data is monthly returns while late hour trading is a daily effect. Hence, the late hour trading effect could be marginal.

The traditionally asset pricing model, known as the Capital Asset Pricing Model (CAPM), uses only one variable to describe the returns of a portfolio or stock with the returns of the market as a whole. Fama & French (1992) used multiple value factors to elaborate the explanatory power of CAPM. They described that size and book-to-market factors with market factors are more powerful to define the return behaviour of various portfolios. Fama & French (1992) proposed 3 factor model for asset pricing by expanding CAPM by adding size risk and value risk factors to the market risk factor.

The Fama-French 3 factor model explains over 90% of the diversified portfolio returns, compared with an average of 70% using CAPM (within sample). The size premium, SMB, is constructed as the difference in returns between the smallest and largest firms. Similarly, HML in each period is the difference in returns between high and low book-to-market firms. This model considers the fact that value and small cap stocks outperform markets on a regular basis. By including these two additional factors, the model adjusts for this outperforming tendency, which is thought to make it a better tool for evaluating managers performance. According to Fama and French, the value stocks out-perform growth stocks. Similarly, small cap stocks tend to out-perform large cap.

Otten & Bams (2002) stated that investment style on the performance of mutual funds can be understood from the Fama-French 3 factor model. They stated that European mutual funds seem to prefer smaller stocks and stock with high book-to-market ratios(value). Carhart (1997) reported that US mutual funds prefer small stocks and stocks with low book-to-market ratios (growth). The important point of the model is that it allows investors to weight their portfolio so that they have greater or lesser exposure to each of the specific risk factors, therefore can target more precisely different levels of expected return.

Bailkowski & Otten (2010) used 4 factor Carhart model to demonstrate that a negative intercept indicate mutual funds do not add value, while a positive intercept indicate mutual funds are able to significantly beat the market. The paper highlighted that mutual fund strategy of buying winners produces significantly positive alpha's while the losers produce significantly negative alpha's. More important that paper also conclude that top funds as well as bottom funds yield positive returns, and this result give us the clue that underperforming fund manager's manipulate returns.

The analysis of properties of returns has a long history of anomalies. Based on the existing literature, it is also found that mutual funds are engaged in a practise of window dressing. Therefore, my dissertation examine monthly returns of US equity mutual funds. To investigate the misreporting by mutual fund, I investigate abnormality in the returns of pooled data across all months from August 2010 to October 2018. The features under investigation are 1) multimodality in the monthly distribution, 2) presence of January effect, 3) presence of seasonality in mutual funds and 4) annual reporting month effect in mutual funds.

The remainder of this dissertation are organised as follows. In Section 2 Data collection, Section 3 considers Methodology adopted for analysis, and Section 4 Concludes the dissertation.

2 Data collection

The dataset of equity type mutual fund monthly returns has been downloaded from a Bloomberg terminal. The sample period is from August 2010 to October 2018. The dataset is from USA of open ended type funds with equity allocation greater than or equal to 70%. There were 2462 funds in this category. Out of 2462 funds, I have eliminated funds with incomplete records over this period. I found that there are 1739 funds with complete records of monthly returns for the period from August 2010 to October 2018. This dataset is the main source for my analysis. To support the evidence of multimodality, I have downloaded monthly returns of Money market (debt) type funds and Close ended equity fund for the same period from August 2010 to October 2018 for comparison.

My main dataset does not include the financial crises of 2008. To further investigate presence of multimodality and January effect during the financial crises; I have also downloaded monthly returns of 1739 mutual funds for the extended period from February 1998 to October 2018. Based on the presence of anomaly in the mutual funds returns, I investigate further to find different features of mutual funds returns. I use Fama-French three factor model to investigate whether these three factors are the features of mutual fund returns. My analysis involves linear factor models for mutual fund returns. I have collected three Fama-French factors that is:

1. the excess return of market,
2. the returns of size and
3. value portfolios,

from Kenneths French's website.

To investigate whether Net Asset Value (NAV) is the feature of mutual funds return, I have downloaded month end NAV of each 1739 mutual funds from Bloomberg Terminal for the period from August 2010 to October 2018. I have also downloaded the annual reporting dates of each mutual fund to investigate whether the reporting month is the feature of mutual fund returns. Here, I assume that mutual funds have not changed their reporting month for the given period. I also investigate seasonality is the feature of mutual fund returns. A statistical summary of monthly returns is:

Name	Particulars/ Type of funds	Period of monthly returns	No. of funds	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	No. of NA's
Monthly returns(equity1)	Equity type mutual fund	August 2010 to October 2018	1739	-48.99	-1.17	1.17	1.00	3.31	31.49	-
Monthly returns(equity2)	Equity type mutual fund	February 1998 to October 2018	752	-67.10	-1.03	3.23	6.06	11.78	88.11	76
NAV	Month end NAV of equity type mutual fund	August 2010 to October 2018	1739	0.97	13.96	19.93	27.13	30.59	113.6	8
Monthly returns(close)	Close ended equity fund	February 1998 to October 2018	101	-86.08	-1.56	3.44	5.97	13.85	114.70	8564
Monthly returns(debt)	Money market (debt) fund	February 1998 to October 2018	1010	-42.42	-0.02	0.69	1.65	2.30	80.68	67651

Table 1: A statistical summary

3 Methodology

The monthly return dataset of mutual funds are a set of continuous data. In many cases it is not possible to specify parametric probability model for my data or it may be too cumbersome to compute the likelihood of the model. A non parametric probability density estimate is suitable for such data which can avoid, the need for such a parametric model. A continuous-valued random variable X has probability density function (pdf), f(x), from which probabilities associated with X can be determined using the accumulative distribution function:

$$\boldsymbol{f(x) = P(X \leq x) = \int_{-\infty}^{x} f(x)dx}.$$

I estimate the pdf f(x) from a sample of observations $x_1, x_2, \ldots, x_n$. There are two common approaches for estimating f(x) referred to as parametric and non-parametric approaches.

Parametric Approach:

The parametric approach for estimating f(x) is to assume the population follow some parametric family of distributions, e.g. $N(\mu, \sigma^2)$, and then to estimate the parameters of the assumed distribution from the data. For example, one could estimate the normal density function f(x) using the estimator:

$$\boldsymbol{\hat{f}(x) = \frac{1}{\sqrt{2\pi}\hat{\sigma}} e^{\left(\frac{x-\hat{\mu}}{2\hat{\sigma}^2}\right)}, x \in R},$$

where $\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} x_i$ and $\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \hat{\mu})^2$.

This approach has various advantages provided the distribution assumption is correct. The main disadvantage of the parametric approach is the lack of flexibility. Each parametric family of distributions imposes restrictions on the shape that f(x) can take. For example, the density function of the normal distribution is symmetrical and bell shaped with exponential tails, and therefore is unsuitable for representing skewed on heavy tailed bimodal densities.

Non-parametric Approach:

The non parametric approach avoids those restrictive assumptions. A well-known non parametric estimator of the pdf is the density histogram. It has the advantage of simplicity but it also has disadvantages, such as lack of continuity and poor tail estimates and sensitivity to binwidth choice.

To construct a density histogram from a continuous variable, the data is split into intervals, called bins. The number of data points in each bin is counted and scaled to ensure it is a proper density. The choice of bins especially the binwidth has a substantial effect on the properties of $\hat{f}(x)$. Note the estimates are the piecewise constant and they are strongly influenced by the choice of binwidth. The bins are typically of equal width i.e. binwidth. 'n' in the total sample size.

1. Define k class intervals $(c_0, c_1], (c_1, c_2], ..., (c_k - 1, c_k]$.
2. Determine the frequency,f_i, of each class i.
3. Calculate the relative frequency (proportion) of each class by dividing the class frequency by the total number in the sample — that is, $\frac{f_i}{n}$.

4. For a density histogram: draw a rectangle for each class with the class interval as the base and the height equal to $h(x) = \frac{f_i}{n(c_i - c_{i-1})}$. As such the total area of all the bins is 1.

A common alternative to estimate the pdf is a Kernel Density Estimator (KDE). Basically, a KDE smoothes each data point $x_1, x_2, \ldots x_n$ using a kernel function from an i.i.d sample, over a small area and then sum all the contribution together to obtain the final density estimate.

The KDE is given by:

$$\hat{f}(x) = \frac{1}{n}\sum_{i=1}^{n} K_h(x - x_i), \text{ and}$$

$$\hat{f}(x) = \frac{1}{nh}\sum_{i=1}^{n} K(\frac{(x - x_i)}{h}),$$

where $K(.)$ is the kernel function, n is the sample size, and the h is the bandwidth. $K(.)$ and h do not depend on x. K_h is called scaled kernel given as:

$$K_h(x) = \frac{1}{n}K(\frac{x}{h}),$$

Generally, the kernel function is a continuous, unimodal, and symmetric density function. To estimate the pdf, I imagine that a scaled normal pdf is centred over each observed data point on the x-axis. The estimator of pdf at a given point is $1/n$ times the sum of the heights of all the normal distribution that cover the point. Figure 1 below shows the weighting functions.

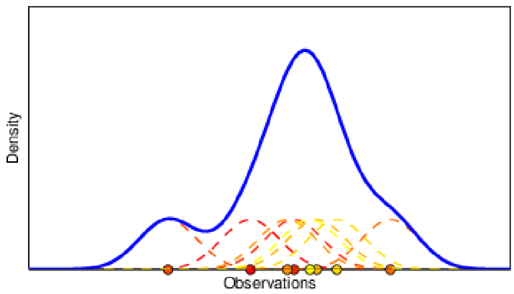


Figure 1: Plot of weighting functions for each point and the combined effect of all weights. Source: www.researchgate.net

A kernel function usually has three key properties:
1. Symmetric about zero
2. $\int K(x)dx = 1$; and
3. $\lim_{x \to -\infty} K(x) = \lim_{x \to +\infty} K(x) = 0$

Most kernel functions are positive; however, higher order kernel functions can be negative.

Using different kernel functions will produce different estimates. Commonly used kernel functions are: 1. Normal or Gaussian, 2. Uniform, 3. Epanechnikov or Quadratic, 4. Triangle

The three forms of kernel functions are:

Gaussian:	$K(x) = \frac{1}{\sqrt{2\pi}} e^{(\frac{x^2}{2})}$
Uniform:	$K(x) = \frac{1}{2} I(-1 \leq x \leq 1)$
Epanechnikov:	$K(x) = \frac{3}{4} max\{1 - x^2, 0\}$
Triangle:	$K(x) = (1 - \lvert x \rvert) \quad \lvert x \rvert \leq 1$

Table 2: Kernel function

Gaussian kernel function is a bell curve shape and it is typically used. It produces smooth density curve while uniform and tiangle do not produce smooth density. Epanechnikov gives an estimate with lowest (asymptotic) mean square error. Figure 2 shows different KDE functions.

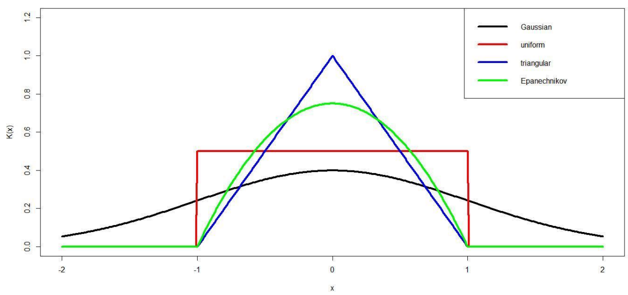


Figure 2: Plot of different KDE function

Histogram provides visual evaluation. I first identify smooth distribution that captures the salient features of the empirical distribution. This smooth distribution serves as reference in the density function.

The most important parameter of a histogram is the binwidth because it controls the tradeoff between presenting a picture with too much details ("undersmoothing") or too little detail ("oversmoothing") with respect to the true distribution.

The optimal binwidth can be chosen to minimise the AMISE (absolute mean integrated square error). Choosing the binwidth is always a trade-off between bias and variance.

Exploratory Data Analysis:

The exploratory data analysis is methodology uses histograms of mutual funds monthly returns to assess whether the underlying densities possess anomalous behaviour. For histogram analysis, the most important parameter is the bin width. Selection of binwidths also depends on objective of the analyser. In this study, we are interested in the number of modes rather than precise location of modes as in spectroscopy.

To select the optimal binwidth, we use of subjective selection method. The subjective selection method in which we experiment by using different binwidth and simply select one that "looks right". The most effective bandwidth selection method is a visual assessment. We (visually) compare different density estimates, based upon a variety of bandwidths, and then choose the bandwidth that corresponds to the (subjective) optimal estimate. The advantage of subjective selection method is human eye takes both horizontal and vertical distances into account when assessing the distance between two curves, while MIAE(mean integrated absolute error) and MISE(mean integrated square error) only measure vertical distance.

The choice of bandwidth for KDE is analogous to choice of bin width for the histogram. Here we assume the underlying distribution is Gaussian when choosing the bin width. The optimal bin width is similar issue to finding the optimal bandwidth.

Bollen and Pool (2009) used the Silverman (1986) method to estimate the bin width. In this method, the bin width is selected for minimum mean squared error between true distribution and the histogram. We have selected bin width as 0.5 on basis subjective selection method while the bin width in the paper of Bollen and Pool (2009) using algorithm in Silverman (1986) is 0.58.

The formula for optimal bandwidth given by Silverman (1986) is:

$$\boldsymbol{h \approx 1.06\sigma n^{\frac{-1}{5}}}$$

where σ is sample standard deviation, n is sample size. For the dataset, we found the bin width 0.9662 with standard deviation 4.0534 for sample size of 1739.

To investigate the prevalence of anomaly in the distribution of monthly returns, I plot month wise histogram of monthly returns of mutual funds from August 2010 to October 2018. The Figure 3 does not show obvious anomalie.

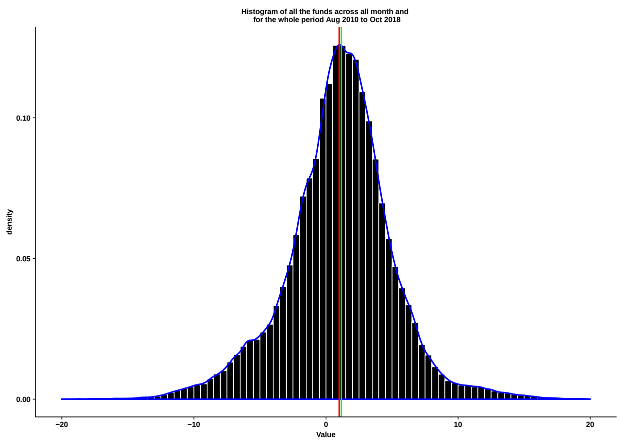


Figure 3: Histogram plot for all the months across all years from August 2010 to October 2018

(In this dissertation, all distribution plots have been generated with sample mean in 'red' line and sample median in 'green' line.)

However, we find subsequent anomalies in the month wise histogram plot of monthly returns of US mutual funds in Figure 4 when compared to histogram plot of Money market (Debt) funds and Close ended funds in Appendix Figure 6.1 and 6.2. The month wise histogram of mutual funds shows different distributions for each month. For few months the distribution is unimodel as expected while for some other months is bimodal or multimodal. Few months distribution have unsual shapes such as July, September.

Now,we further investigate the presence of different features in the mutual funds, which are prevalent in hedge funds as reported by researchers, is also prevalent in mutual funds. These features are 1) multimodality in the monthly distribution, 2) presence of January effect, 3) presence of seasonality in mutual funds and 4) annual reporting month effect in mutual funds.

1) Multimodality in the monthly distribution

We investigate multimodality in the monthly distribution. To find the presence of multimodality in the distribution, we plot month-wise histogram for the period of August 2010 to October 2018. We plot month wise plot of histogram of monthly returns of all funds across all years. The Figure 4 below shows the month wise plot of histogram. Note for all histogram plot in this dissertation we discard outliers, if any. Outliers are observations that fall outside the overall pattern.

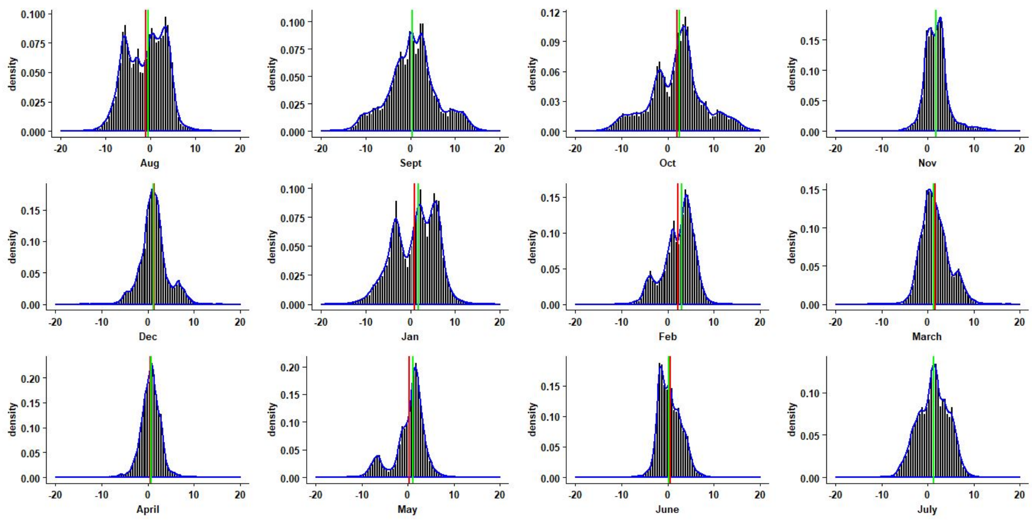


Figure 4: Histogram plot of each the month across all years from August 2010 to October 2018

From Figure 4, we observe following type of distribution for each for months:

Month	Distribution	mean and median
April, November	Unimodal	mean = median > 0
June	Skewed unimodal	$mean \geq median \geq 0$
March	Asymmetric bimodal	$median > mean > 0$
May, October, December	Asymmetric bimodal	$median < mean < 0$, for December $0 < mean = median$
January	Triimodal	$median > mean > 0$
February	Asymmetric trimodal	$median > mean > 0$
August	multimodal	$mean < median < 0$
September, July	undefined	for July $median > mean > 0$, for September $median = mean \leq 0$

Table 3: Type of month-wise distribution

From the Figure 4, we conclude that multimodality is a feature of mutual fund returns. We compare the monthly distribution with that of Money market (debt) funds, and Close ended equity funds. We find that monthly distribution of Money market (debt) funds and Close ended equity funds are unimodal, while that of mutual funds show multimodality certain months particularly January and August. This is evidence for the presence of abnormality in the return distribution of mutual funds.

The pick of the histogram has the highest density that means majority of the mutual funds has monthly return at the pick. We observe that most of the histogram has pick above zero, mean<median, and median is positive. The monthly distributions have slightly moved away from zero to positive side.

Now, we study the image plot of densities month wise across all years to see the density flow per month across all years. This image plot enables us to observe multimodality in details in a particular month such as January and August. Figure 5 shows Image plot for the each month for all years.

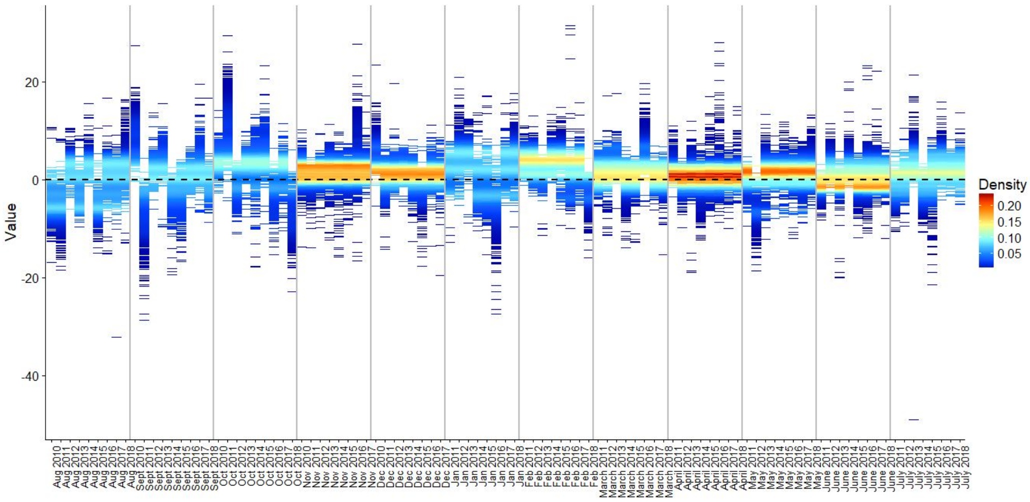


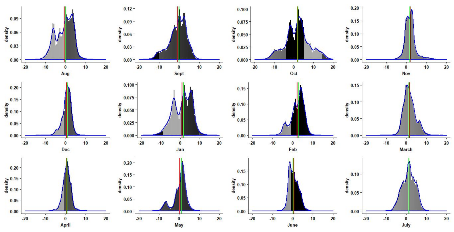
Figure 5: Imageplot of densities of returns monthwise for whole period from August 2010 to October 2018

(Detailed figure of tile plot for each quater in Annexure Figure 6.4 to 6.7)

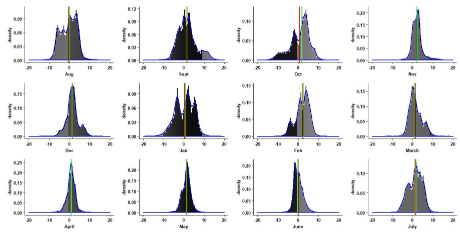
Here,we plot mutual fund returns per month per year with density as colour. It is interesting to note that the image plot of mutual fund returns has mean level above zero. We found that most of the returns are positive with densities concentrated at 0.10 and above (white colour). Further, the months of November, December, April and May have large number of positive returns just above zero followed by March and July. We observe that large number of densities is concentrated just above zero in November, December and large number of funds report in December followed by October and September. The December reporting is followed by a mysterious January multimodality effect. The January effect is a hypothesis about the increase in stock prices lead by decrease in prices in December. This effect seems to affect small caps than large or medium cap because they are less liquid. January effect suggests that market is inefficient.

Further, we investigate 1739 mutual funds for the longer period from February 1998 to October 2018. The reason to study anomaly for the longer period is to find the changes in the distribution due to 2008 financial crisis. Secondly, we want to observe the features are consistent for the longer period. Figure 6.3 in the Appendix shows the month wise plot of histogram. Most of the month histogram is right skewed with fat tail on positive side. Further mean and median of returns are above zero for all the months. We find bimodality distribution for the Janaury month. The multimodality in the August month has disappeared. No remarkable changes in the distribution for the month of September as the stock market crashed in the September 2008.

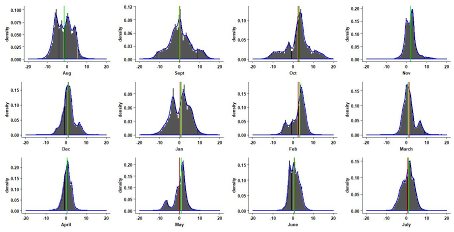
Now, we are interested to find whether year is the feature of mutual funds returns. The reason for us to investigate the yearly effect is to look for major changes in the monthly distribution due to unusual trading in stock market in a year. For example, 2010 Flash crash during which Dow Jones Industrial Average suffered worst intra-day point loss in May . We have plot month wise histogram by removing a year at time from the dataset. The dataset consist of 9 years monthly returns from 2010 to 2018. Figure 6 shows 9 plot of month-wise histogram for the dataset after removing a year each time. It is found that there is not substantial change in the histogram. This is evidence that year is not ariving the multimodal features of mutual fund returns.



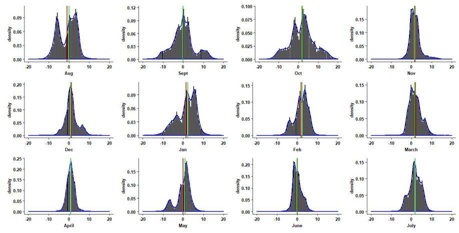
(a) Excluded year 2010



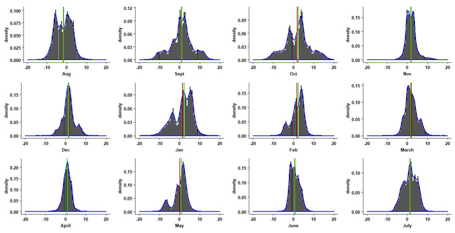
(b) Excluded year 2011



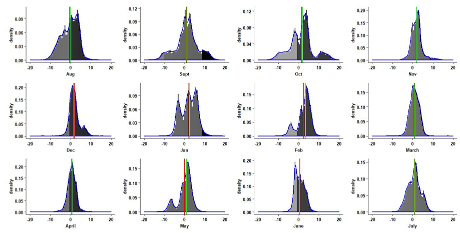
(c) Excluded year 2012



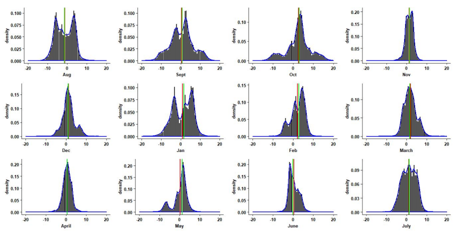
(d) Excluded year 2013



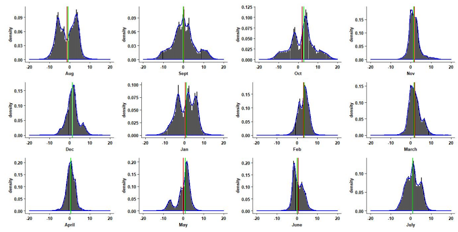
(e) Excluded year 2014



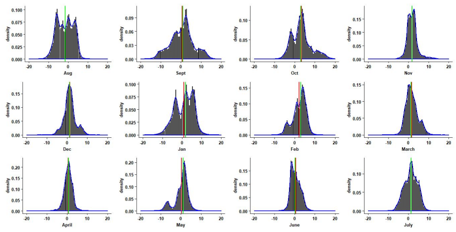
(f) Excluded year 2015



(g) Excluded year 2016



(h) Excluded year 2017

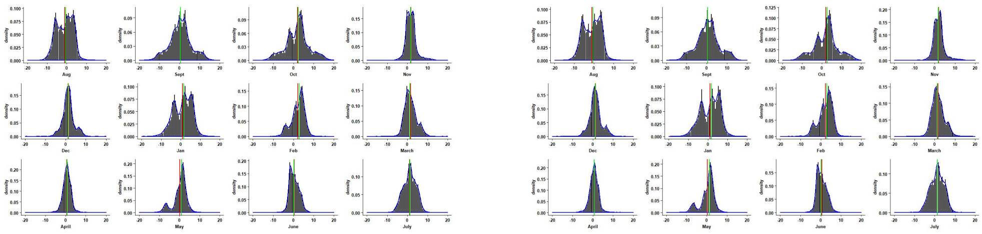


(i) Excluded year 2018

Figure 6: Histogram plot of returns excluding one year at a time

(Detailed figure in Annexure 6.8 to 6.16.)

Now, we study the feature of presence of anomaly in the monthly returns on basis of month-end Net asset value (NAV) of fund. We separate the fund on the basis of average of month-end NAV of each fund is above and below media. On separating the dataset on basis of size, we plot the month wise histogram for both the dataset. Figure 7 shows month wise histogram plot of mutual fund of small and large size.



(a) Returns less than NAV

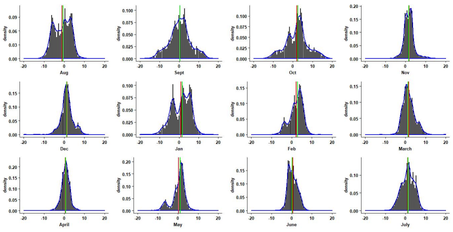
(b) Returns more than NAV

Figure 7: Histogram plot of each the month across all years for dataset with NAV less and more than median

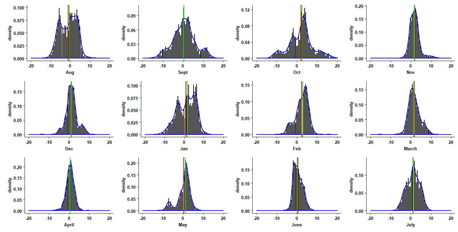
(Detailed figure in Annexure Figure 6.17 and 6.18.)

There is no substantial different in the histogram. Hence we conclude that size (in terms of NAV) of the funds is not ariving the multimodal features of mutual fund returns.

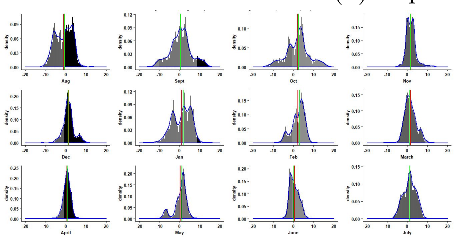
Now, we investigate whether the reporting month of the mutual fund is the feature of mutual funds returns. We found the highest number of funds report their annual financial statement in the month of December, followed by October and September. I found that out of 1739 funds, 367 funds report in December followed by 361 funds report in October and 247 funds report in September. It is important to note that only 68 funds reported in November. We separate the dataset based on mutual funds reporting on decending order of reporting months i.e. December, October and September. Figure 8 shows month wise histogram of funds reporting in December, October and September.



(a) Reporting in December



(b) Reporting in September



(c) Reporting in October

Figure 8: Histogram plot of mutual funds reporting in December, September and October

Detailed figure in Annexure Figure 6.19,6.20,6.21.

We found no substantial differences in the histogram sets depending on the months when the annual financial statement is reported. Hence, we conclude that the reporting month of the mutual fund is not ariving the multimodal features of mutual fund returns.

Now, we implement the Fama-French three factor model to investigate how the style of investment impact its features of mutual fund returns. Fama-French modified the original CAPM with two additional risk factors: size risk and value risk. For a fund 'I' we fit the Fama-French three factor model.

Fama French model:

The Fama-French 3 factor model is given by:

$$\boldsymbol{R_{it} - R_{ft} = \alpha + b_i(R_{mt} - R_{ft}) + s_i SMB_t + h_i HML_t + \epsilon_{it}}$$

where

$\boldsymbol{R_{it}}$ = the return of fund I in the month t

$\boldsymbol{R_{ft}}$ = the return of a one month T-bill in month t

$\boldsymbol{R_{m}}$ = the return on the local equity benchmark in month t

$\boldsymbol{SMB}$ = the difference in return between a small cap portfolio and a large cap portfolio

$\boldsymbol{HML}$ = the difference in return between a portfolio of high (book-to-market ration) minus a portfolio of low (book-to-market ratio)

$\boldsymbol{\alpha}$ = intercept of the model, also called the Jensen Alpha, which is interpreted as a measure of out- or under performance relative to the used market proxy. It is the abnormal returns' intercept of security characteristic line.

$\boldsymbol{R_{it} - R_{ft}}$ = the excess fund return

$\boldsymbol{R_{mt} - R_{ft}}$ = the value of weighted excess return on the market portfolio

$\boldsymbol{\epsilon_{it}}$ = an error term

In this study, we are analysing monthly returns of mutual funds hence t is 1. The coefficients $\boldsymbol{b_i, s_i, h_i}$ are the betas of the stock on each of the three factors, often called the factor loadings. If these factors fully explain asset returns, the intercept of the equation should be zero.

$\boldsymbol{b_i, s_i, h_i}$ can be determined by linear regression and can take negative values as well as positive values. We find loading factors by Ordinary Least Squares regression (OLS) of monthly returns with Fama-French 3 factor data for period August 2010 to October 2018.

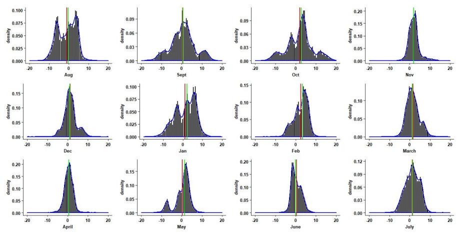
We separate the dataset of monthly returns on basis of 2 classes for each loading factor and for positive and negative value of intercept alpha. The following table shows the statistical summary of loading factors:

Betas	Beta of excess market	Beta of SMB	Beta of HML	Alpha	Intercept of model
Nos funds above Median	869	869	869	Nos of funds with Positive intercept	398
Nos funds below Median	870	870	870	Nos of funds with Negative intercept	1341
Total	1739	1739	1739	Total	1739
Median	0.9988	0.0717	0.0068	Median	-0.1017
Mean	0.9785	0.197	-0.0002	Mean	-0.1211
Min	-0.5090	-0.514	-1.107	Min	-1.9980
Max	2.2920	1.326	1.76	Max	0.6637

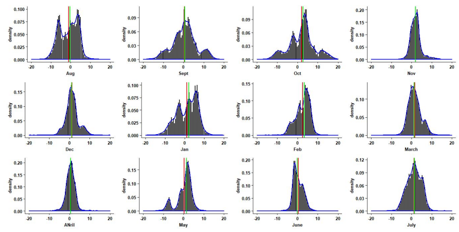
Based on the above data, we separated the monthly returns based on intercept and coefficients. We have 4 variable out of which 3 coefficients are (1) beta of excess returns, (2) beta of SMB, (3) beta of HML and one intercept 'alpha'. In case of intercept, we separate monthly returns with positive intercept and negative intercept. After separating the dataset, we plot month wise histogram for mutual fund returns with positive and negative alpha.

To investigate effects of Fama-French factor, we select one coefficient at a time and divide monthly returns into two separate group for the betas above median and betas below median. After separating the monthly returns,we once again plot the month wise histogram for each dataset across all years to examine investment style is the feature of monthly returns of mutual funds.

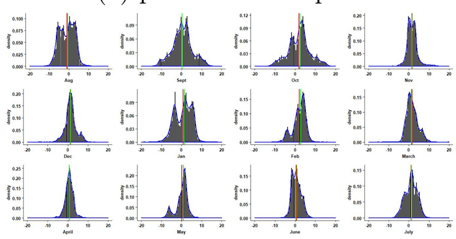
The number of mutual funds with positive intercept is 398 and negative intercept is 1341. This implies that most of the mutual funds are overvalued. From the Figure 9 plot of month wise histogram, we have not found any substantial changes in the month wise plot of histogram for either Fama-French factor or intercept. Thus, we conclude that style of investment or abnormal returns due to intercept is not ariving the multimodal feature of mutual fund returns.



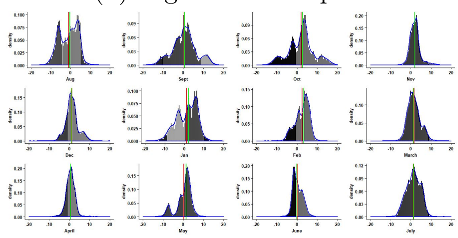
(a) positive intercept $\boldsymbol{\alpha}$



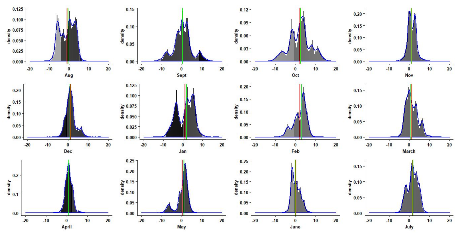
(b) negative intercept $\boldsymbol{\alpha}$



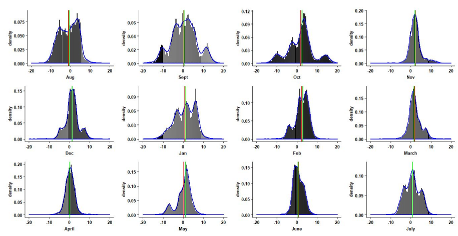
(c) returns below median of beta of excess return



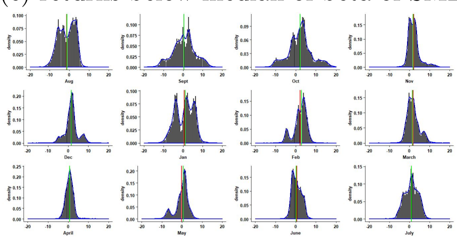
(d) returns above median of beta of excess return



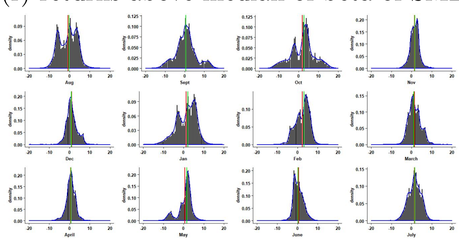
(e) returns below median of beta of SMB



(f) returns above median of beta of SMB



(g) returns below median of beta of HML



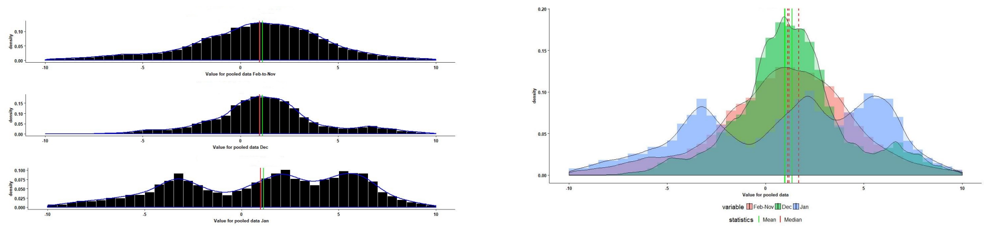
(h) returns above median of beta of HML

Figure 9: Histogram plot of mutual funds reporting above and below median of coefficients

(Detailed figure in Annexure figure 6.22 to 6.29)

2) Presence of January effect

One of the prominent monthly anomaly tested in the world market is the January effect which is also related to tax-loss selling hypothesis. We investigate the dataset for prevalence of January effect. We separate out three set of data of monthly returns of funds into 1) January, 2) December, and 3) February to November. Figure 10 shows that the distribution for the dataset of December and February to November are unimodal however the plot of January is multimodal. It seems to provide support that there exists a January effect.



(a) Separate plot of January, December and February to November

(b) Super embosing plot of January, December and February to November

Figure 10: Histogram of January, December and February to November

From Figure 4, The histogram of January shows a multimodal distribution. The trimodal distribution is the evident of three distinct groups are present in the distribution. From Figure 10, we find that the January month shows higher variability than December and remaining dataset from February to November. It is evidence that there is presence January effect in the distribution as seen in Figure 11.

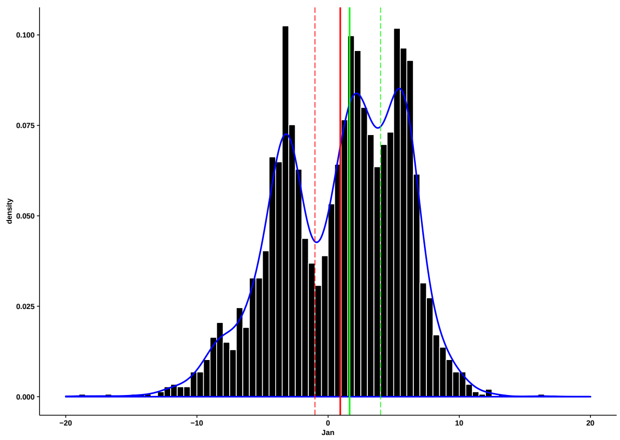


Figure 11: Histogram of January month across period from August 2010 to October 2018

We further investigate the type of funds responsible for the positive and negative returns in the month of January. The histogram of January has unusual modes on the negative and positive side of histogram. We separate the funds reporting below the trough on negative side of histogram and above the trough on positive side of histogram. Figure 11 shows the trifurcation of histogram to find exact number of funds on either side of histogram.

We found there are 258 funds that report negative returns below the lower trough and 262 funds that report positive returns above the upper trough. Further, there are 258 funds which are common in lower trough and upper trough. These same number of funds are reporting both negative and positive returns.

Summary of type of funds found below the lower trough and above the upper trough:

Type of funds as per information given by promoters/fund managers of each fund	Below the lower trough	Above the upper trough
Small cap investments (small, mocro, medium)	87	87
Large cap investments	55	54
Investment in both large cap and small cap / specialist fund	21	21
Investment Status not known	99	96

Summary of all the 1739 funds:

Type of funds as per information given by promoters/fund managers of each fund for 1739 mutual funds	Nos.
Small cap investments (small, mocro, medium)	682
Large cap investments	275
Investment in both large cap and small cap / specialist fund	136
Investment Status not known	646

The above figures give us the rough idea about the investment strategy of each mutual fund. From the summary of mutual funds found below the lower trough and above the upper trough, it is not clear which type of investment style is associated with January effect.

Now, we check the mean-median of monthly returns for February to November, December and January.

Month	Mean	Median
February to November	0.9728	1.1309
December	1.3495	1.1985
January	0.9836	1.6896

When the above results are studied along with the distribution, we find that mean of December month is substantially higher than rest of months with unimodal distribution while median of January is substantially higher than rest of months with multimodal distribution.

When closely observed the image plot (Annexure 6.5) of January,we find that majority of mutual funds report positive and negative returns. This is the reason for trimodal distribution of the January monthly returns. As per existing research papers on hedge funds by Agarwal, Daniel & Naik (2011) concluded that December returns are sustantially higher than January. Further, the research paper on mutual fund by Hyunk-Suk Choi (2015) also concluded that equity fund experiences high and low net cash flow in January and in December respectively. Hence window dressing is the feature of some funds. However it cannot be confirmed that the January effect in the dataset is associated to investment style of the funds i.e micro, small, medium, large or specialist fund. From Fama-French model, we have found no evidence that the January effect is the feature of investment style as per Figure 10 and 11. The multimodality of distribution in the month of January is not fund specific.

However the effect weakens for the period from February 1998 to October 2018. Figure 6.3 in Annexure shows that the January effect has weakened as the trimodality vanishes.

3) Presence of seasonality in mutual funds

Seasonality refers to some features or patterns in time series repeat every year. To investigate presence of seasonality, I plot the month wise histogram by removing a year at a time. My dataset of monthly returns for the period of August 2010 to October 2018 has 9 years; I will have 9 plots of month-wise histogram after excluding returns of particular year each time to observe the changes.

Some literature has pointed out presence of seasonality. The sell in May effect in equities is well documented but it has weakened as it has become part of public information as documented by Degenhardt & Auer (2018), Bouman (2001). During the 'winter' months stock markets perform much better than summer months. Hyunk-Suk Choi (2015) established presence of seasonality in cash flows to US domestic mutual funds.

To demonstrate that seasonality effect is the feature of mutual fund, we plot month-wise histogram of monthly returns by removing dataset of each year at a time to see the variation in histogram. In Figure 6, I find no substantial different in each month's histogram plot after excluding a year each time. Hence, it is evident that mutual funds of equity type with equity proportion greater than or equal to 70% have seasonality behaviour.

4) Annual reporting month effect in mutual funds.

Now, we would investigate annual reporting month effect in mutual funds. Here, we define annual reporting month effect as those month when mutual funds submit their annual report, followed by subsequent months which are subject to changes in returns that signal good or bad times for investment.

To confirm the presence of annual reporting month effect in mutual funds, we plot the histogram relative to the reporting month. We have collected the latest annual reporting month of each fund from Bloomberg Terminal. We assume that mutual funds have not altered reporting months during the period 2010 to 2018. We rearrange the dataset of monthly returns of each fund such that the monthly returns starts from reporting month upto end of financial year of that particular fund. We ran this process for all the funds. After rearranging all the 1739 funds starting with reporting month as the zeroth month, we now plot the histogram starting from zeroth month of reporting to 11th month from reporting month for each fund.

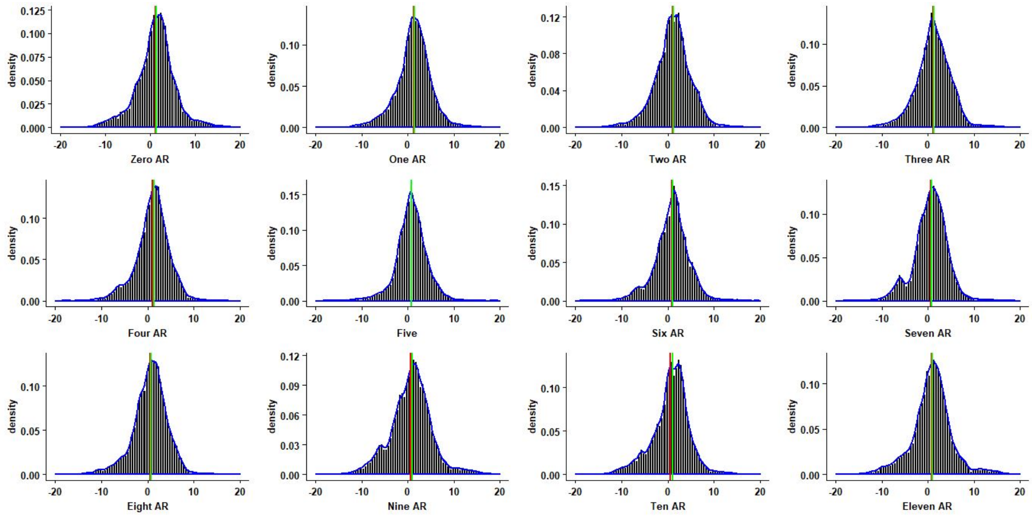


Figure 12: Histogram plot of return based on distance from reporting month

Figure 12 shows that unimodal histogram for each monthly distance from reporting month. The multimodality feature which was found earlier in the month-wise plot of histogram for the same period has disappeared. This results provide the clue that most of the mutual funds have single mode. In other way, we understand that they have similar behaviour with respect to distance from reporting month. This indicates a single value is dominant among others. With the appearance of unimodal histogram, the dataset has aligned in position with respect to reporting months. Instead of strong multimodality in distribution, rather it has weaker.

We conclude that annual reporting months is not ariving the multimodal feature of mutual funds.

4 Conclusion

This dissertation gives persistent and robust results in finding abnormality in the return distributions of US mutual funds for the period August 2010 to October 2018. I further investigate the abnormality by way of finding features in returns. This is the basis for my research. The overall findings of the study showed some noteworthy outcomes for investors.

Examining the potential anomaly in US mutual funds, I found abnormality in the month-wise distribution of returns. Particularly, I found multimodality in the distribution for the month of January and August. I have also found that same funds are reporting positive and negative returns in the distribution of January. This is in consistent with previous literature study on mutual fund by Hyunk-Suk Choi(2015) and that on hedge fund by Agarwal, Daniel & Naik (2011) and call for further research in this area. Here, it is important to note that some of the mutual funds are involved in window dressing. While trying to understand different features of mutual fund returns, I came across important findings listed below:

1. I have found that year is not ariving the multimodal feature of mutual fund returns. I have not found major changes in the distribution with inclusion or exclusion of financial crisis of 2008. There was huge fall in stock market globally on September 2008. However such fall has not affected the distribution of September. In fact for longer period, all the month wise distributions were unimodal except January which is bimodal.
2. When size of the mutual funds in the form of NAV was examined on the return distribution, I find that size is not ariving the multimodal feature of mutual fund returns. When reporting months were considered for analysis, the top three reporting months were December, October and September are favourites among mutual fund managers to submit annual reports. Based on the distribution of those mutual funds reporting in these three months did not show major differences in the distribution. I did not find evidence for reporting months as the multimodal feature of mutual fund returns.
3. The Fama - French three factor model provide a good criterion to study the return behaviour. I found no evidence of ariving that Fama-French factors as the multimodal feature of mutual fund returns. In other way, I can say that the investment style is not ariving the multimodal feature of mutual fund returns. However, I found the distribution of the factor

SMB to be right skewed, which supports the argument that mutual funds in US prefer small firm or 'small firm effects'.

4. I found multimodality in the distribution of January. However, I have not found any evidence for the cause of such effects. Initially, it was assumed that capital size is the feature of January effect in mutual funds. However, I could not find any particular type of mutual fund responsible for January effect.

5. When examined for the impact of seasonality effect on the return behaviour of mutual funds; I found seasonality as the feature of mutual funds.

6. The noteworthy finding of the study is that the reporting month and the succeeding months is not ariving the multimodal feature of mutual fund returns. The return distribution turns unimodal once the data are aligned on the basis of reporting months from 0th month to 11th month from annual reporting month. I call this annual report effect

The anomaly in form of abnormality in the distribution is evident that mutual fund returns are consistent to seasonality effect. As Thomas Kuhn has pointed out in his book *The Structure of Scientific Revolutions*, scientific advances typically occur when an accepted paradigm does not explain repeated observation of anomalies. Thus, the reported results called for further investigation for the reason being (1) trimodal distribution found in January and August remains unknown, (2) seasonality effect and (3) annual report effect.

5 Reference

Agarwal, Vikas, Gerald D. Gay, and Leng Ling, “Window dressing in mutual funds,” The Review of Financial Studies, 2014, 27(11), 3133–3170.

Agarwal, Vikas, Naveen D. Daniel, and Narayan N. Naik, “Do Hedge Funds Manage Their Reported Returns?,” The Review of Financial Studies, Oct. 2011, V(24), No. 10, 3281-3320.

Bialkowski, Jedrzej, and Roger Otten, “Emerging Market Mutual Fund Performance: Evidence for Poland,” University of Canterbury Research Repository, 2010.

Bollen, Nicolas P.B. and Veronika K. Pool, “Do Hedge Fund Manager Misreport Returns? Evidence from the Pooled Distribution,” The Journal of Finance, Oct. 2009, LXIV(5), 2257-2288.

Bouman, Sven, “The Halloween Indicator, ‘Sell in May and Go Away’: Another Puzzle,” Eramus University Rotterdam, July 2001, 1-34.

Brown, Keith C., W. V. Harlow, and Laura T. Starks, “Of tournaments and temptations: An analysis of managerial incentives in the mutual fund industry,” Journal of Finance, 1996, 51, 85-110.

Carhart, Mark M., “On Persistence of Mutual Fund Performance,” The Journal of Finance, 1997, 52(1), 57-82.

Chevalier, Judith, and Glenn Ellison, “Risk taking by mutual funds as a response to incentives,” Journal of Political Economy, 1997, 105, 1167-1200.

Cox, R. Don and Ken Johnston, “The January Effect is not Driven by Tax Loss Selling,” The Journal of Investing, 1998, 7(4), 105-111.

Degenhardt, Thomas, Benjamin R Auer, “The ‘Sell in May’ effect: A review and new empirical evidence,” The North American Journal of Economics and Finance, 2018, V43, 169-205.

Fama, F. Eugene and Kenneth R. French, “The Cross-Section of Expected Stock Returns,” The Journal of Finance, June 1992, XLVII(2), 427-465.

Fama, F. Eugene and Kenneth R. French, “Common risk factors in the returns on stocks and bond,” The Journal of Financial Economics, 1993, 33, 3-56.

Grinblatt, Mark, and Sheridan Titman, “Mutual fund performance: An analysis of quarterly portfolio holdings,” Journal of Business, 1989, 62, 394-416.

Gruber, Martin J., “Another Puzzle: The Growth in Actively Managed Mutual Funds,” The Journal of Finance, July 1996, V(51), 783-810.

Gultekin, Mustafa N. and N. Bulent Gultekin, “Stock Market Seasonality: International Evidence,” Journal of Financial Economics, 1983, 12, 469-481.

Haugen, Robert A., and Josef Lakonishok, “The incredible January effect: The stock market’s unsolved mystery, Dow Jones-Irwin, Homewoof, Ill,”1988.

Holmes, William. R, “A practical guide to the Probability Density Approximation (PDA) with improved implementation and error characterization,” Journal of Mathematical Psychology, 2015, 68-69, 13-24.

Huang, Jennifer, Kelsey D. Wei, and Hong Yan, “Participation costs and the sensitivity of fund flows to past performance,” Journal of Finance, 2007, 62, 1273-1311.

Hyun-Suk Choi, “Seasonality in Mutual Funds Flows,” The Journal of applied Business research, March/April 2015, v31(2), 715-726.

Jorge Perez, “Are Un-Registered Hedge Funds More Likely to Misreport Returns?,” University of Albany, State university of New York, 2014, 5.

Kacperczyk, Marcin, Clemens Sialm, and Lu Zheng, “Unobserved Actions of Mutual Funds,” The Review of Financial Studies, Nov. 2008, 2379-2416.

Keim, Donald B., “Size Related Anomalies and Stock Return Seasonality,” Journal of Financial Economics, June 1983, 13-22.

Landauskas, Mantas, “Modelling of stock prices by the Markov Chain Monte Carlo Method,” Intellectual Economics, 2011, V(5), No. 2(10), 244-256.

Ledl, Thomas, “Kernel Density Estimation: Theory and Application in Discriminant Analysis,” Australian Journal of Statistics, 2004, V(33),No.3, 267-279.

Mehta, Kiran and Ramesh Chander, “Application of Fama and French Three Factor Model and Stock Return Behavior in Indian Capital Market,” Asia-Pacific Business Review, Dec. 2010, 12-17

Otten, Roger, and Dennis Bams, “European Mutual Fund Performance,” European Financial Management, 2002, 8(1), 75-101.

Ritter, Jay R., and Navin Chopra, “Portfolio rebalancing and the turn-of-the-year effect,” Journal of Finance, 1989, 44, 149-166.

Roll, Richard, “Vas ist Das? The Turn-of-the-Year Effect and the Return Premia of Small Firms,” Journal of Portfolio Management, Winter 1983, 18-28.

Sirri, Erik R., and Peter Tufano, “Costly search and mutual fund flows,” Journal of Finance, 1998, 53, 1589-1622.

Thaler, Richard H. “The January Effect,” Economic Perspectives, Summer 1987, V(1), 197-201.

Tinic, Sneha M. and Richard R. West, “Risk and Return: January and the Rest of the Year,” Journal of Financial Economics, 1984, 13, 561-574.

Zitzewitz, Eirc, “How widespread is late hour trading in Mutual Funds?,” Stanford Graduate School of Business, 2004 (working paper).

Zucchini, Walter, “Applied smoothing techniques,” Temple University, Philadelphia, Pennsylvania, Oct. 2003, 1-19.

6 Appendix

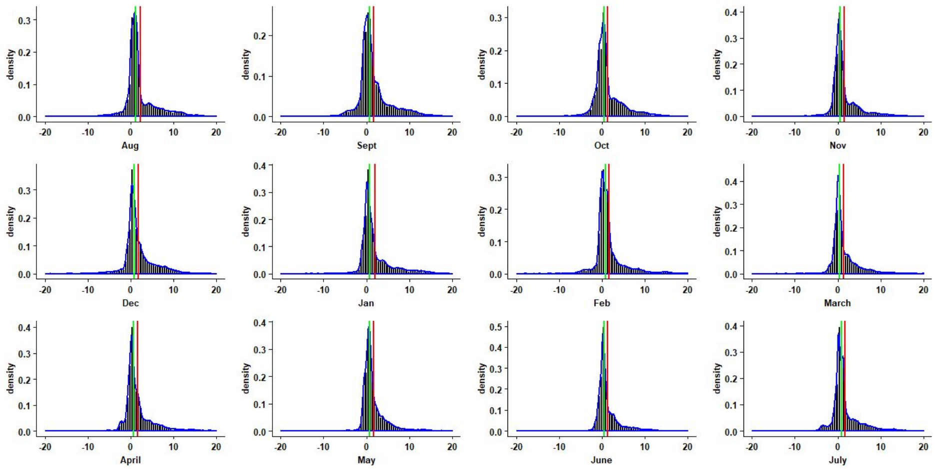


Figure 6.1: Histogram plot of Money Market(Debt) fund month wise across all years from February 1998 to October 2018

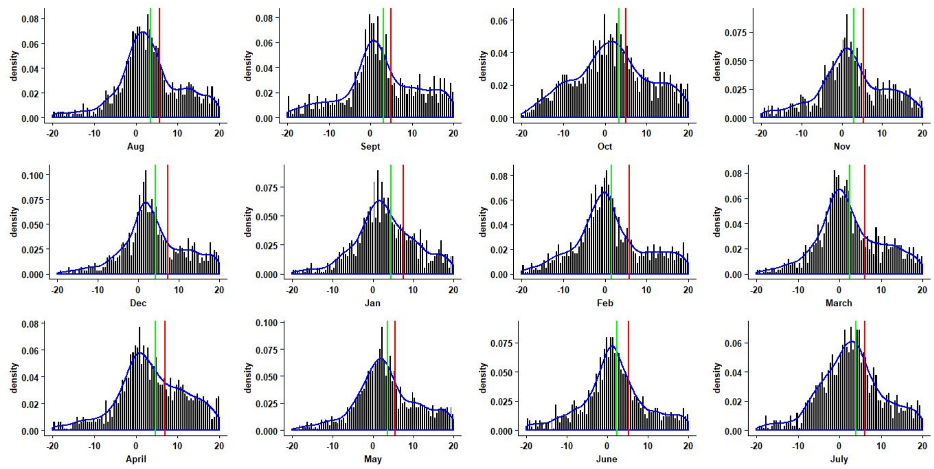


Figure 6.2: Histogram plot of Close ended fund month wise across all years from February 1998 to October 2018

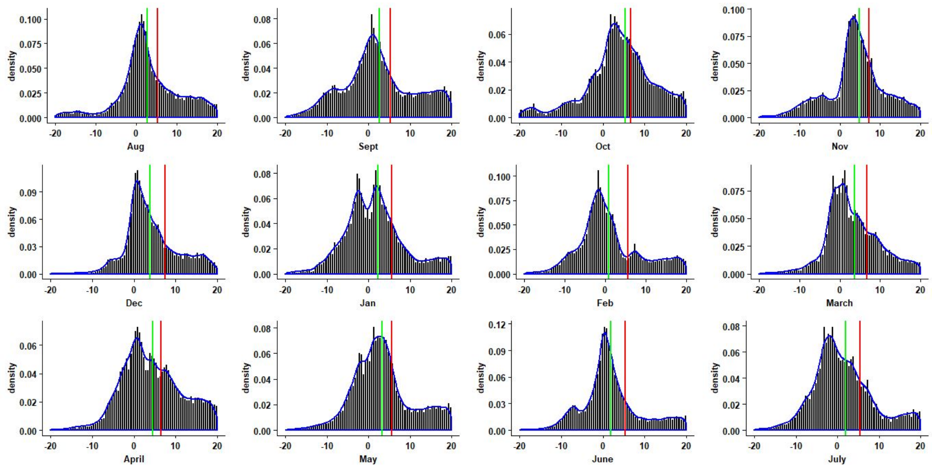


Figure 6.3: Histogram plot of Mutual funds month wise across all years from February 1998 to October 2018

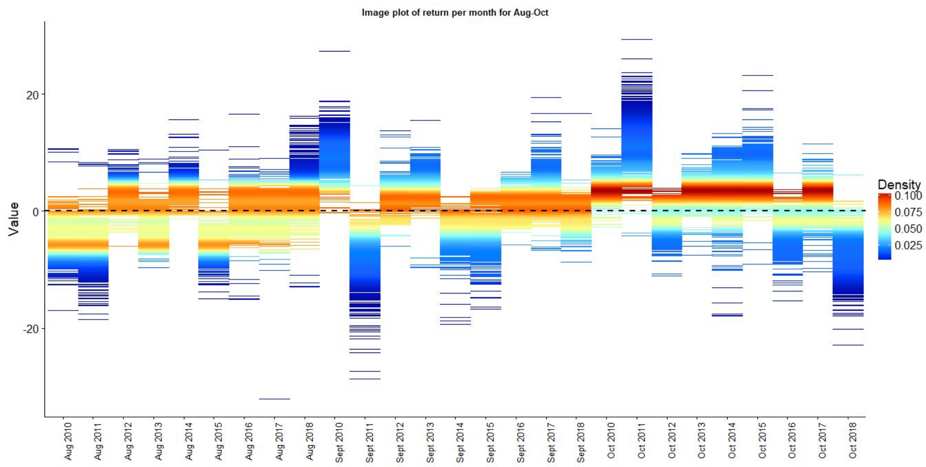


Figure 6.4: Tileplot of densities of returns monthwise for period from August to October across all year 2010-2018

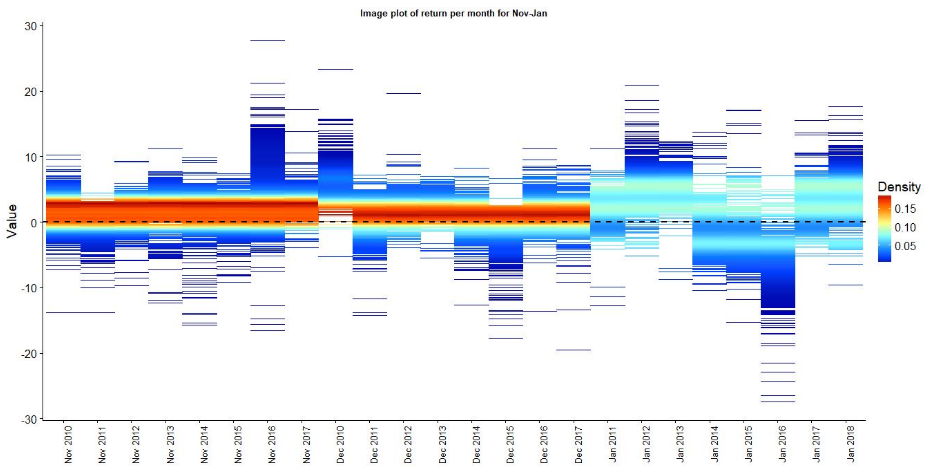


Figure 6.5: Tileplot of densities of returns monthwise for period from November to January across all year 2010-2018

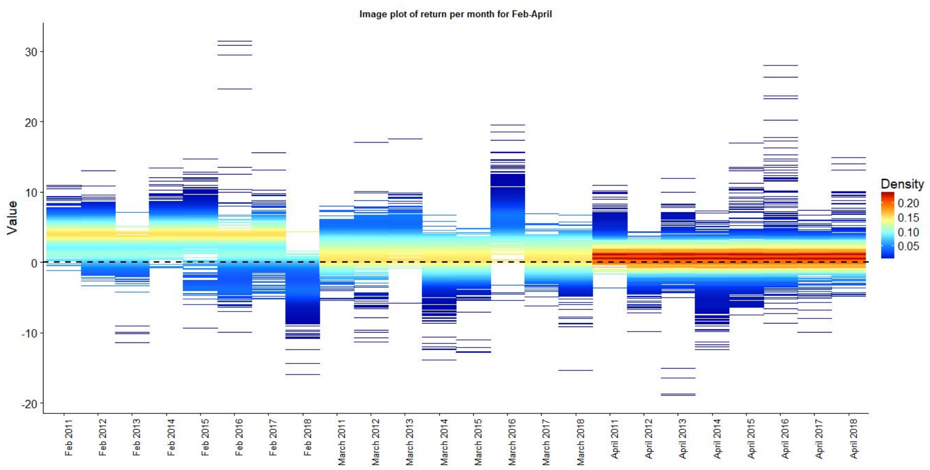


Figure 6.6: Tileplot of densities of returns monthwise for period from February to April across all year 2010-2018

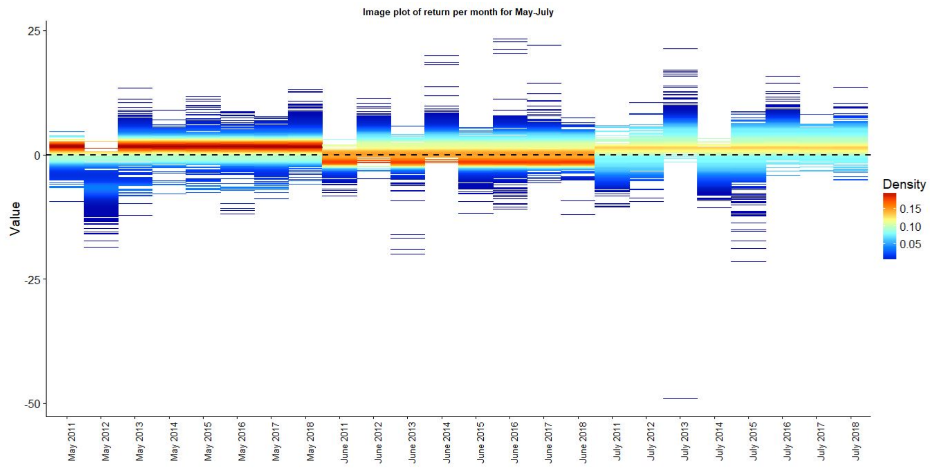


Figure 6.7: Tileplot of densities of returns monthwise for period from May to June across all year 2010-2018

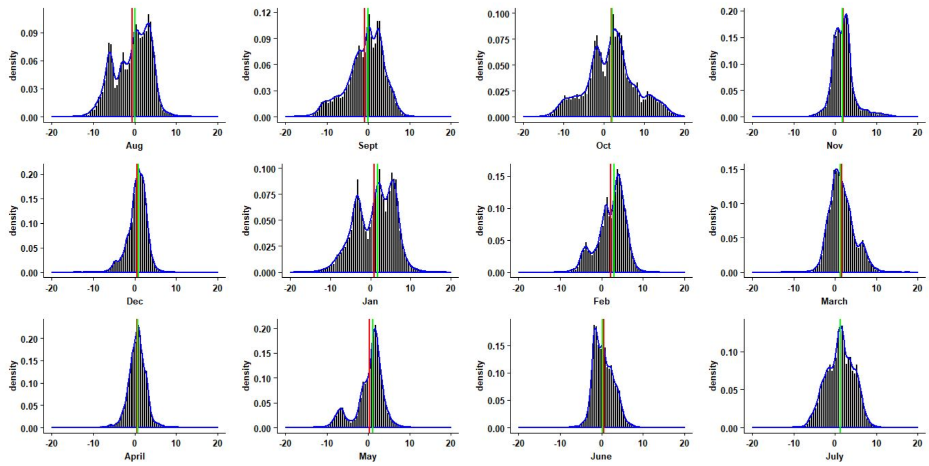


Figure 6.8: Histogram plot of each the month across all years from August 2010 to October 2018 excluding year 2010

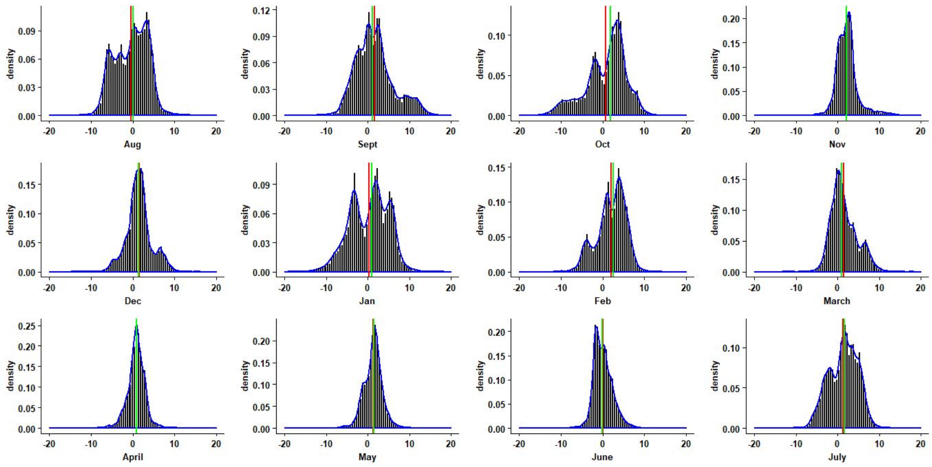


Figure 6.9: Histogram plot of each the month across all years from August 2010 to October 2018 excluding year 2011

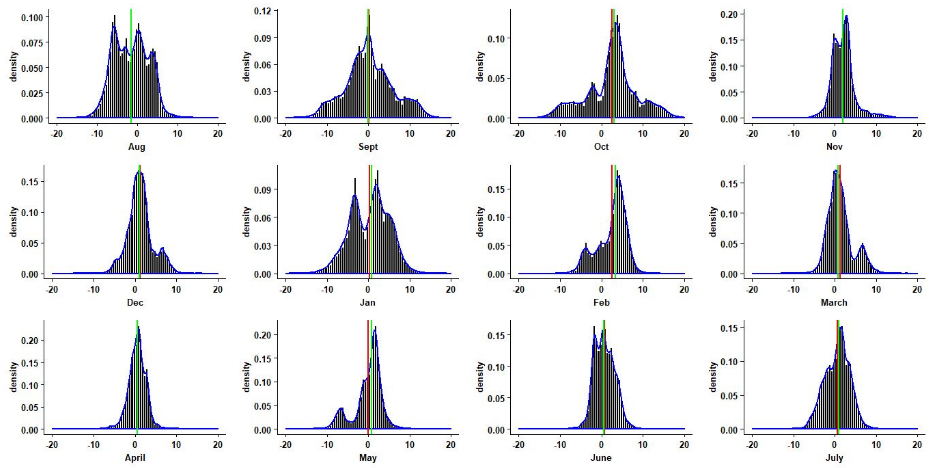


Figure 6.10: Histogram plot of each the month across all years from August 2010 to October 2018 excluding year 2012

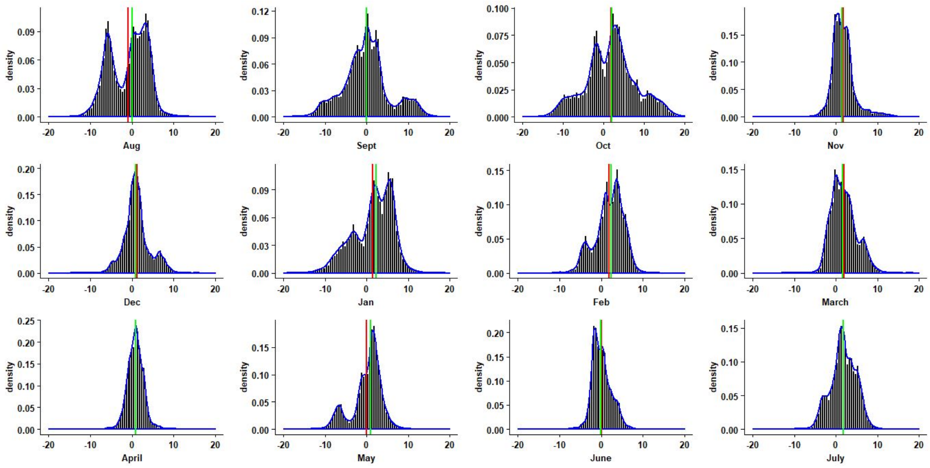


Figure 6.11: Histogram plot of each the month across all years from August 2010 to October 2018 excluding year 2013

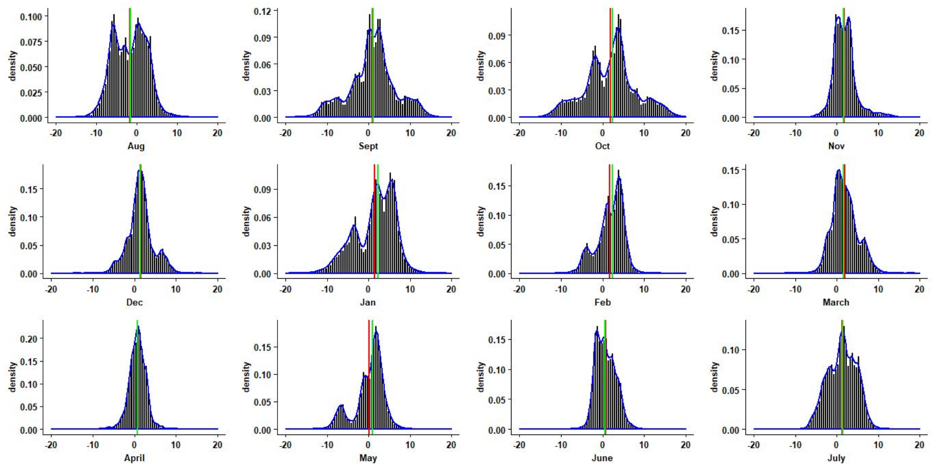


Figure 6.12: Histogram plot of each the month across all years from August 2010 to October 2018 excluding year 2014

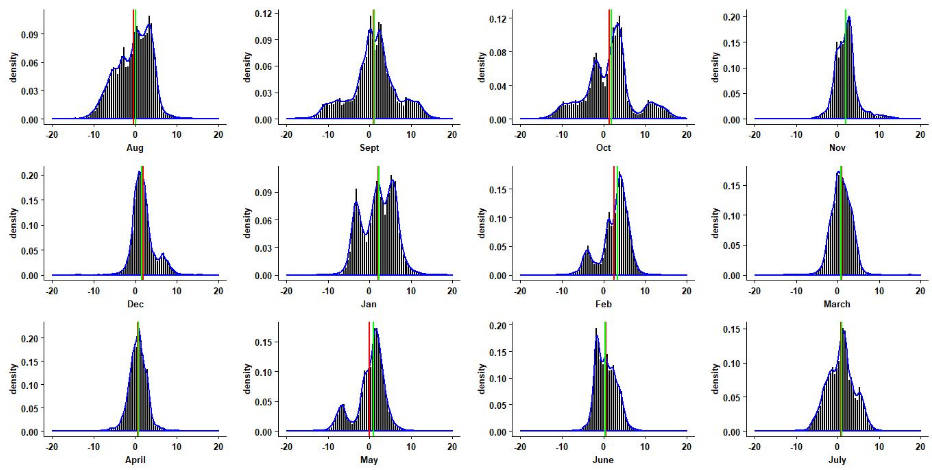


Figure 6.13: Histogram plot of each the month across all years from August 2010 to October 2018 excluding year 2015

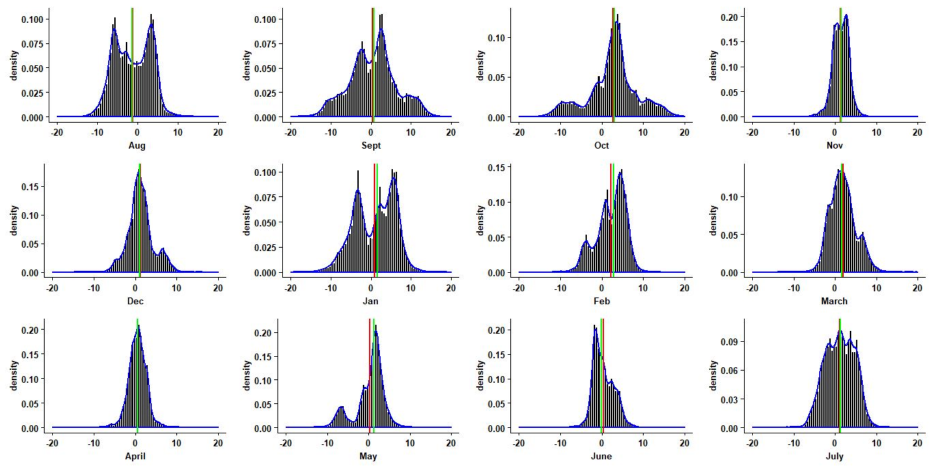


Figure 6.14: Histogram plot of each the month across all years from August 2010 to October 2018 excluding year 2016

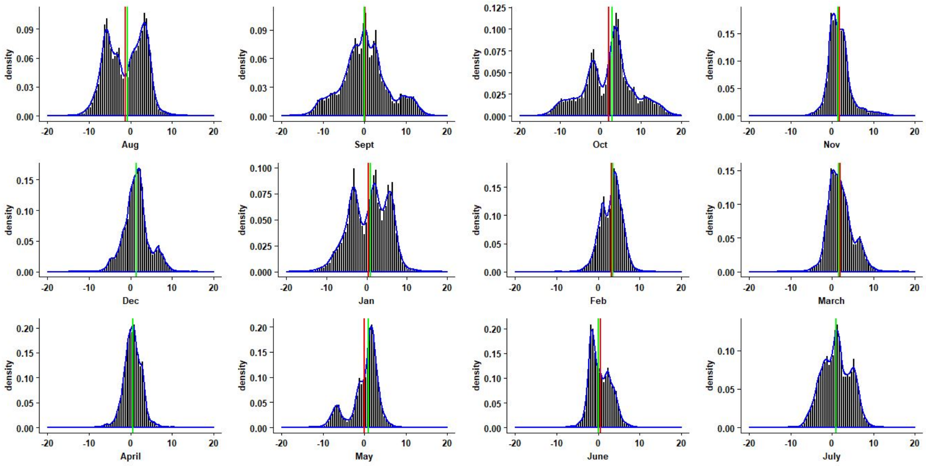


Figure 6.15: Histogram plot of each the month across all years from August 2010 to October 2018 excluding year 2017

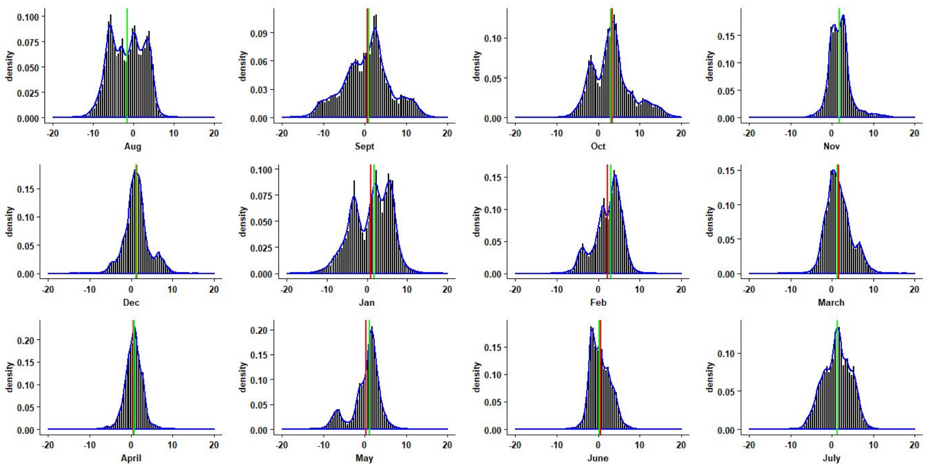


Figure 6.16: Histogram plot of each the month across all years from August 2010 to October 2018 excluding year 2018

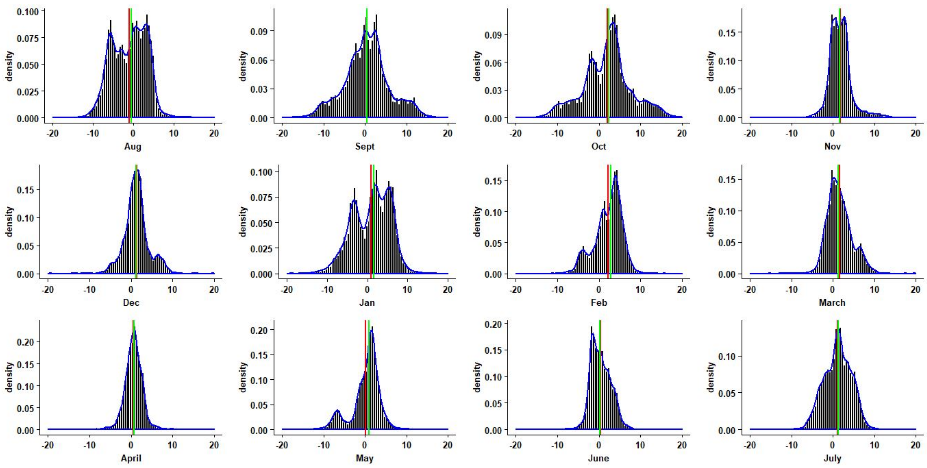


Figure 6.17: Histogram plot of each the month across all years for returns lower than NAV of each fund

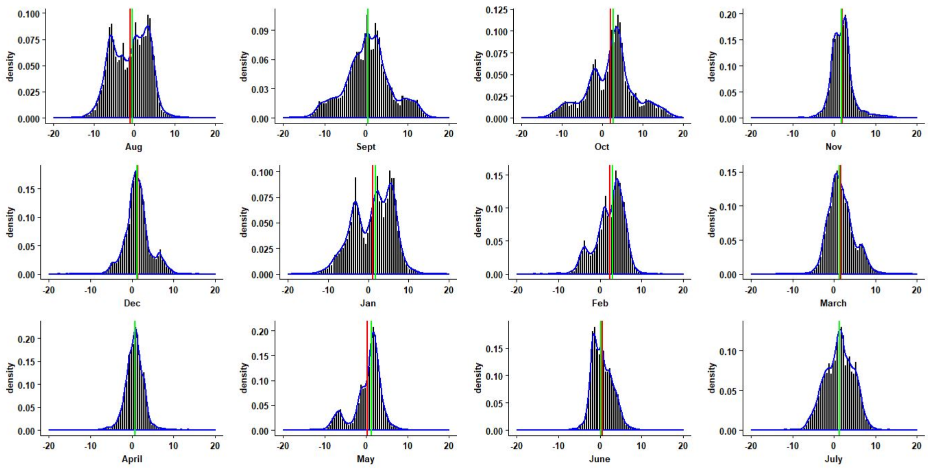


Figure 6.18: Histogram plot of each the month across all years for returns higher than NAV of each fund

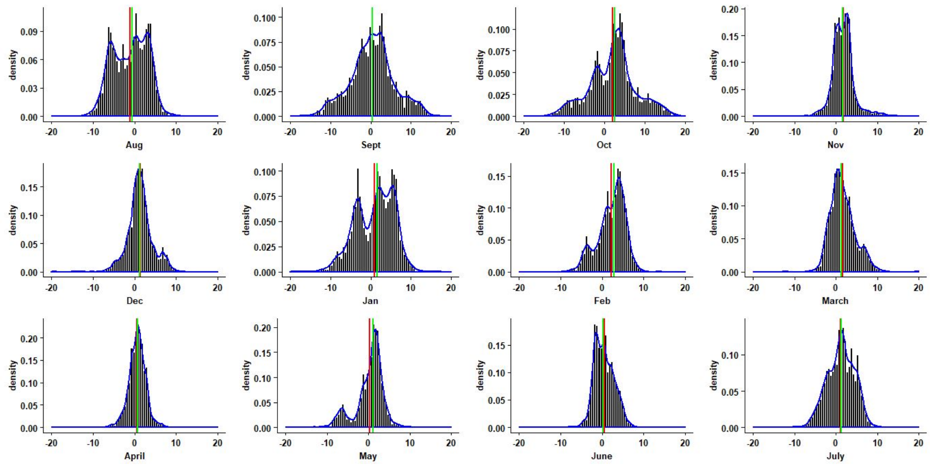


Figure 6.19: Histogram plot of each the month across all years for funds reporting in December

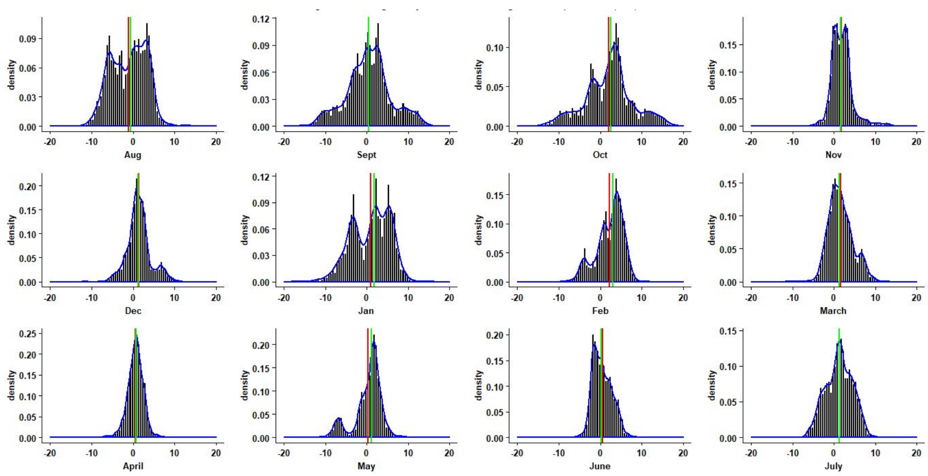


Figure 6.20: Histogram plot of each the month across all years for funds reporting in October

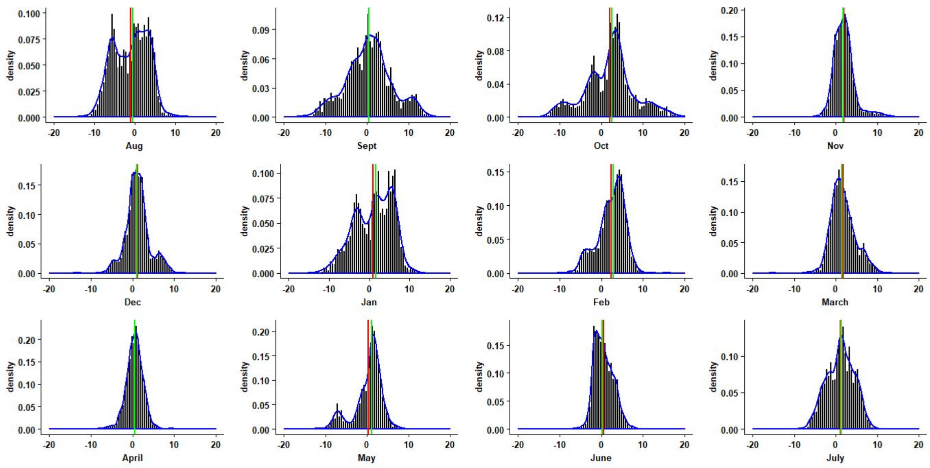


Figure 6.21: Histogram plot of each the month across all years for funds reporting in September

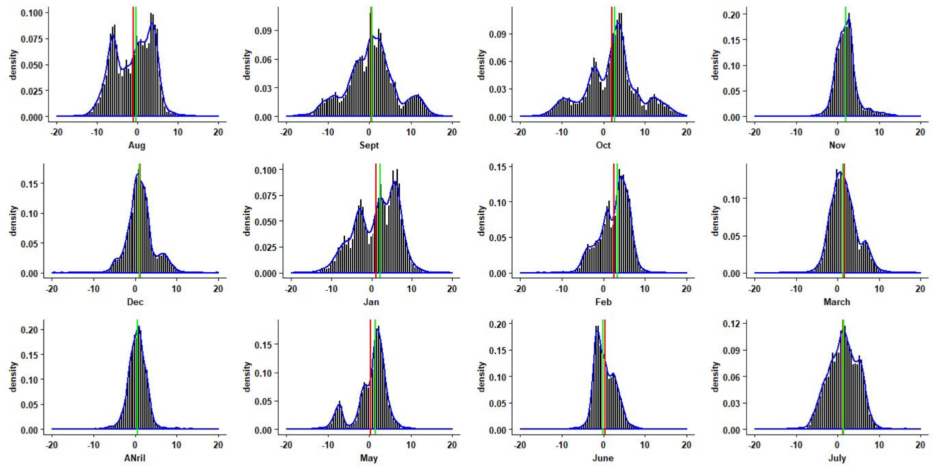


Figure 6.22: Histogram plot of each month across all years of returns for funds with negative alpha

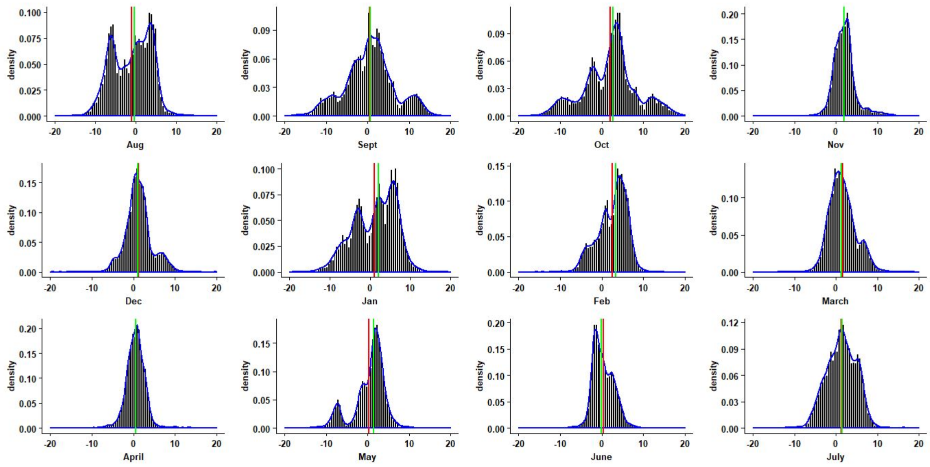


Figure 6.23: Histogram plot of each month across all years of returns for funds with positive alpha

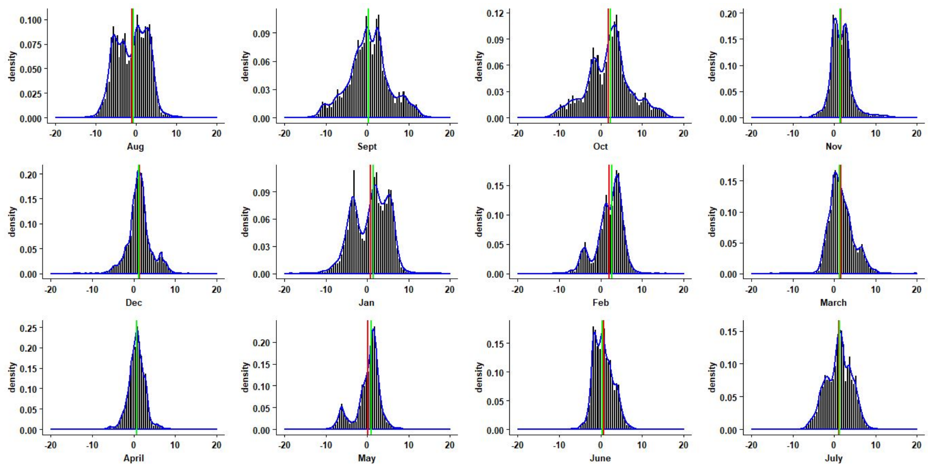


Figure 6.24: Histogram plot of each month across all years for returns of funds beta less than median of market returns

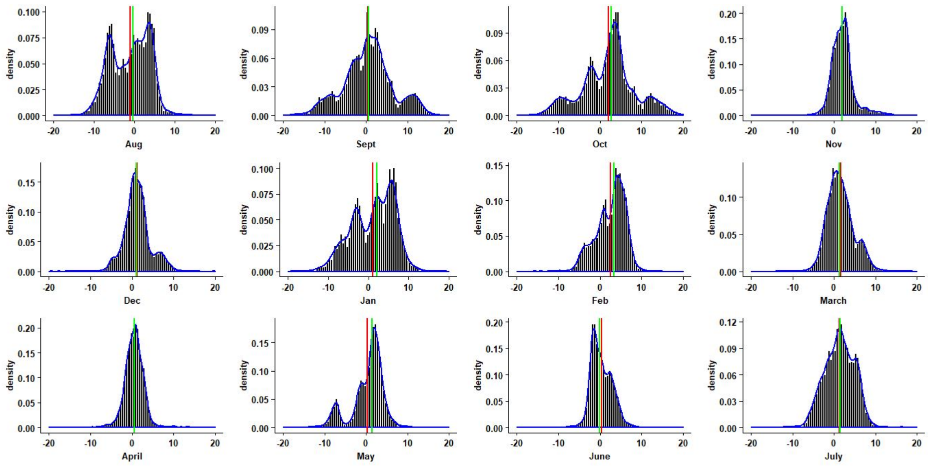


Figure 6.25: Histogram plot of each month across all years for returns of funds beta more than median of market returns

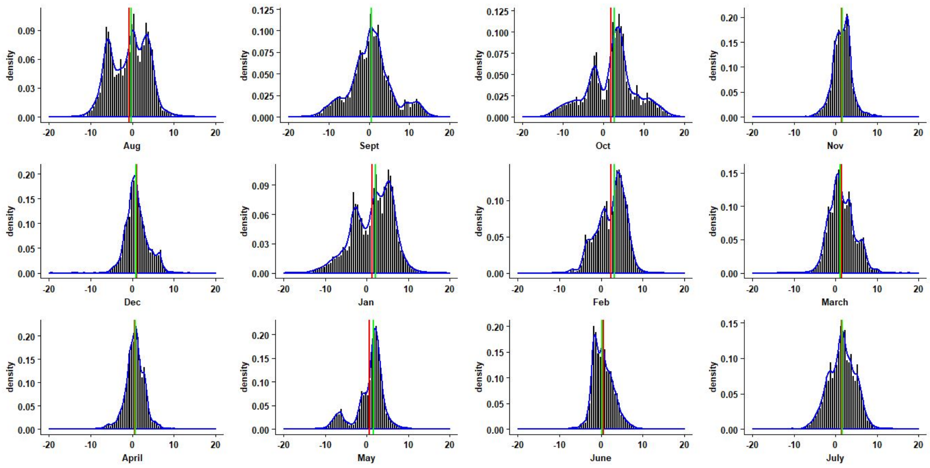


Figure 6.26: Histogram plot of each month across all years for returns of funds beta less than median of HML

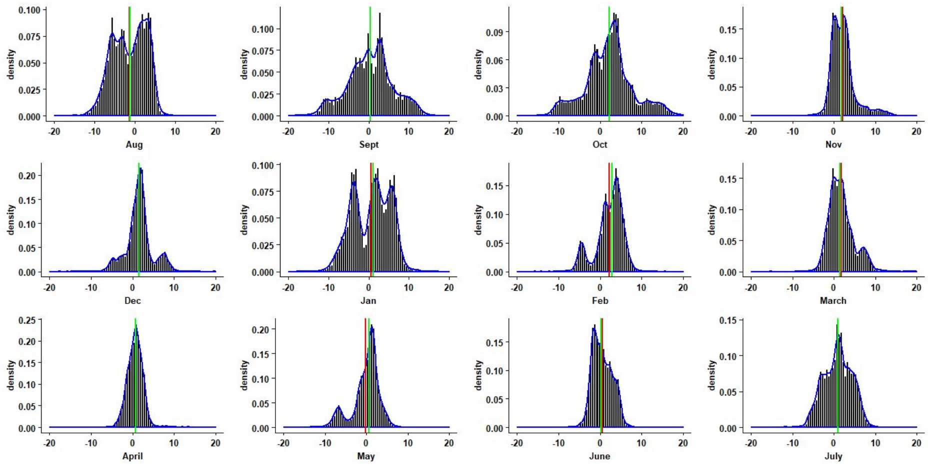


Figure 6.27: Histogram plot of each month across all years for returns of funds beta more than median of HML

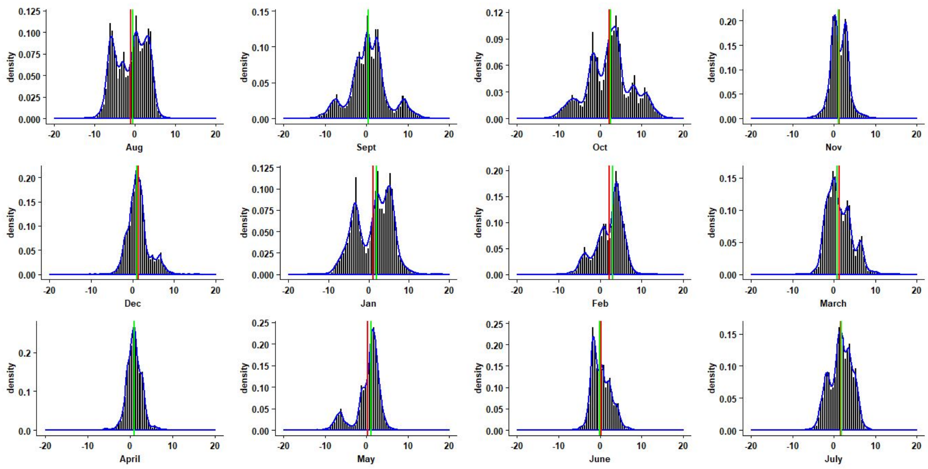


Figure 6.28: Histogram plot of each month across all years for returns of funds beta less than median of SMB

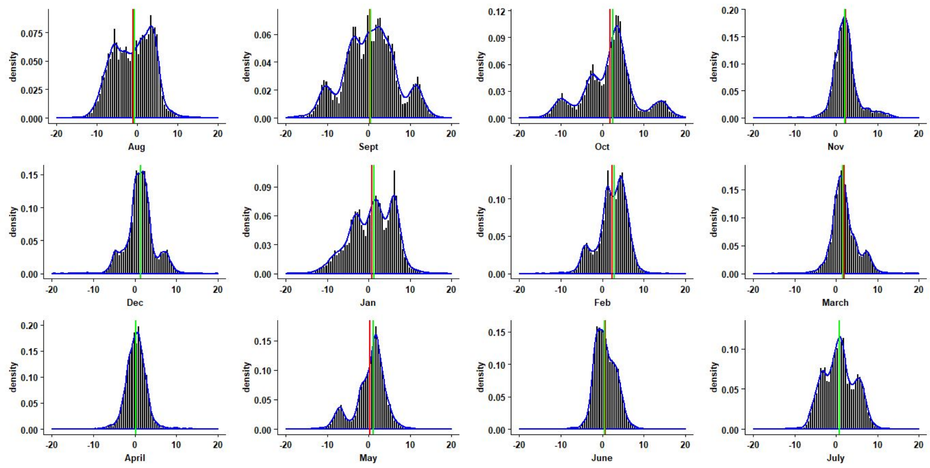


Figure 6.29: Histogram plot of each month across all years for returns of funds beta more than median of SMB