Review of Finance (2013) 17: pp. 1743–1785
doi:10.1093/rof/rfs035
Advance Access publication: December 5, 2012

Are Monthly Seasonals Real? A Three Century Perspective

CHERRY Y. ZHANG and BEN JACOBSEN*
Massey University

Abstract. Over 300 years of UK stock returns reveal that well-known monthly seasonals are sample specific. For instance, the January effect only emerges around 1830. Most months have had their 50 years of fame, showing the importance of long time series to safeguard against sample selection bias, noise, and data snooping. The overall conclusion is that monthly seasonals might simply be in the eye of the beholder.

JEL Classification: G10, G14

1. Introduction

Had stock markets been a field of academic study early in the nineteenth century, our predecessors would have wondered about the significantly positive August and December effects and asked themselves why stocks performed so poorly in October. Researchers in the early 1900s pondering a century of stock market returns might have tried to explain the significantly negative July and August effects.

To what extent are seasonal stock market anomalies real? In their seminal study, Lakonishok and Smidt (1988) prescribe long and new data series as the best medicine against data snooping, noise, and 'boredom' (selection bias). They confirm many daily anomalies, like the Turn of the Month effect and the Turn of the Week effect, in their extended sample of 90 years of the Dow Jones market index. As they point out at a monthly

*This article has benefited from comments from an anonymous referee, Sven Bouman, Burton Hollifield (the editor), Walter Torous, and seminar participants at the 2011 Financial Management Association Asian meeting in Queenstown, New Zealand, and the 2011 New Zealand Financial Colloquium and presentations at several universities. The usual disclaimer applies.

level, however, they add little new data and even a 90-year sample offers no remedy using monthly frequency data[1]:

Monthly data provides a good illustration of Black's (1986) point about the difficulty of testing hypotheses with noisy data. It is quite possible that some month is indeed unique, but even with 90 years of data the standard deviation of the mean monthly return is very high (around 0.5 percent). Therefore, unless the unique month outperforms other months by more than 1 percent, it would not be identified as a special month.(Lakonishok and Smidt, 1988, p.422)

While new data sets of long time series of stock returns are becoming available, no article has used these data to verify whether monthly seasonals are real or are chimeras. This article fills that gap by looking at over 300 years of monthly data on the UK stock market, starting in 1693. We use these UK data as it is the longest time series available and also provides us with a relatively fresh new data set, as they have been less mined than the data from the USA.

Contrary to the Lakonishok and Smidt (1988) results, where their longer sample period confirmed well-known daily effects, our longer series sheds new light on many monthly calendar anomalies. Many months significantly under- or outperform over the full period and in subperiods, but few have done so persistently throughout the ages. This suggests that monthly calendar anomalies change over time, or that these anomalies do not exist. Whether or not anomalies exist seems to depend strongly on the chosen sample period and sample length. We illustrate this using the full sample but also sample lengths of a 100 years (close to the 90 years suggested in the quote above) and 50 years (as proxy for the smaller sample sizes used by most other studies).

Whether or not these anomalies exist also depends on how one weighs the statistical evidence. If one requires an anomaly to be statistically significant and with consistent signs in all subperiods of reasonable length and across different estimation methods (ordinary least squares (OLS), Generalized AutoRegressive Conditional Heteroskedasticity (GARCH), and robust regressions), there may be no monthly anomalies. If one feels that Lakonishok and Smidt's argument above has some merit—that we need at least 90 years or more to establish reasonable confidence bounds—we should rely on the

[1] Increasing the interval of observation does not answer this question either, as Merton (1980, p. 365) points out: "Accuracy of the (expected return) estimator . . . depends only upon the total length of the observation period . . . nothing is gained in term of accuracy of the expected return estimate by choosing finer observations intervals for the returns . . . ."

longer samples or full sample evidence only. Based on the full sample only, the evidence points to seasonal effects in 4 months (significantly positive in January and December and negative in July and October) and a significant Halloween or Sell in May effect. These effects are significant and robust across estimation methods in the full sample. Changing the weights one uses to evaluate the statistical evidence leads to a different combination of anomalies. In short, it seems safe to say that whether or not these anomalies do exist, is in the eye of the beholder, and depends strongly on the sample used and which criteria are applied.

No month—including January—significantly outperforms the market persistently in all our 50- and 100-year subsamples, although December comes close, only exhibiting below average returns in the first half of the twentieth century. In the first 150 years, instead of being the best performing month, January is significantly worse than average. Before 1830, there is a strong positive December effect, which weakens as the January effect emerges. Only July almost consistently underperforms in our full sample and in all of the 50- and 100-year subsamples. However, not even in the subsamples of a 100 years does it always underperform significantly. Moreover, if we use 50-year rolling window regressions, we find periods with positive July returns as well. The 50-year rolling regressions nicely illustrate the point of Lakonishok and Smidt (1988) that these sample sizes are too small for reasonable statistical inference. Unfortunately, even the 100-year subsamples do not seem to provide unambiguous evidence either.

This long monthly series also allows us to test the persistence of the Sell-in-May effect, or the Halloween effect (Bouman and Jacobsen, 2002), which is the notion that winter returns (November through April) are substantially higher than summer returns (May through October). Bouman and Jacobsen (2002) find this anomaly present in 36 out of 37 countries. Many studies have confirmed the existence of this Halloween effect in stock returns.[2] Bouman and Jacobsen (2002) also present evidence of the effect in the same UK data as we use over the full 300-year period starting from 1693. Nonetheless, they leave open the possibility that this anomaly may also have varied over time. We consider that possibility here. We confirm their result for the full sample. But we cannot confirm it has always been significantly present in the subsamples as well. Measured over 100-year intervals, it

[2] For instance, Swinkels and Van Vliet (2010) find a US equity premium over the sample period of 1963–2008 of 7.2%, if there is a Halloween effect and a Turn of the Month effect, and a negative risk premium of −2.8% in all other cases. We discuss more studies in Section 2.

is always positive but not always significant. Measured over 50 years, the effect tends not to be significant in the first 100 years and in the beginning of the twentieth century, it is sometimes negative (although not significantly so). Again, the anomaly may be in the eye of the beholder. However, if one believes, it does exist that it has dramatically increased in strength since the 1950s.

Our focus on the long-term history of UK data is especially interesting, as the UK is the home of the market wisdom *Sell in May and go away*. Popular wisdom suggests that the effect originated from the English upper class spending winter months in London, but spending summer away from the stock market on their estates in the country: an extended version of summer vacations as we know them today.[3] Thus, if the Sell in May anomaly should be significantly present in one country over a long period, one would expect it to be the UK.

A number of studies have made profound contributions in making high-quality historical time series data available, allowing others to test and revisit current findings in the literature. For instance, Goetzmann, Ibbotson, and Peng (2001) construct monthly stock price and total return indices from over 600 individual stocks on the New York Stock Exchange (NYSE) starting from 1815 and running to 1925. Wilson and Jones (2002) improve the monthly S&P stock price index from 1871 to 1999 making it a more consistent broad index. For the UK, Grossman (2002) provides an annual price index with broader coverage of the market to the standard index for the period from 1870 to 1913 and Acheson *et al*. (2009) present a monthly index of total returns for the UK stock market for 1825–1870.

Historical data are used to examine the robustness of current empirical findings and economic theories. For instance, Goetzmann and Ibbotson (2005) document the historical equity premium of the US market back to 1792 and find a relatively stable real rate of returns over the past two centuries. Using stock prices of three big companies traded in both the London and Amsterdam stock markets, Neal (1987) shows that both markets are informationally efficient, with high levels of integration between them. Harrison (1998) examines the distribution and higher moments of Amsterdam and London stock returns in the eighteenth

[3] To give an example: "Historically, the summer fall was caused by farmers selling and sowing their crops and rich investors swanning off to enjoy Ascot, The Derby, Wimbledon, Henley and Cowes. Modern investors jet off to the Med, where they cannot find copies of their pink papers and senior fund managers soak up the sun on Caribbean cruises leaving their nervous second-in-commands in charge" (*The Evening Standard*, May 26, 1999).

century. Brown and Easton (1989) test whether weak form efficiency holds for the 3% Consols in the London market for the period from 1821 to 1860. Brown Jr., Burdekin, and Weidenmier (2006) study the volatility of 3% Consols in the London market from 1792 to 1959, and infer that political stability might be an important explanation for the dramatic decline in volatility during the Pax Britanica period. Jorion and Goetzmann (1999) report the equity premia for 39 countries and the potential diversification benefits of the countries over the period from 1921 to 1996. Goetzmann, Li, and Rouwenhorst (2005) investigate the benefits of international diversification from 1850 onwards and find that the benefit of global investing varies over time. Grossman and Shore (2006) reveal that size and long-term reversal anomalies are not present in the UK market from 1870 to 1913, and that the period only exhibits weak evidence of a value effect. With a 90-year daily time series index from 1897 to 1986, Lakonishok and Smidt (1988) confirm the persistence of many daily anomalies, including the turn of the week, turn of the month, turn of the year, and the holiday effect in the US market. Using data back to 1871, Jones, Pearce, and Wilson (1987) show that the January effect is present long before income taxes in the USA, which goes against the tax-loss selling hypothesis. Similarly, Choudhry (2001) also reports evidence of the January effect in both the USA and the UK in the period from 1870 to 1913.

Others study long time series data to increase the power of the test where small sample inference could potentially bias the results of empirical findings. For example, Shiller (1989) examines the co-movements of stock prices and dividends between the UK and US markets from 1919 to 1987. Goetzmann (1993b) finds evidence of mean reversion of long-term stock returns using 300 years of UK data and 200 years of US data. Goetzmann (1993a) shows a strong positive relation between art demand and the stock market over the period of 1715–1986. Lundblad (2007) confirms the positive relation between the market risk premium and expected volatility using US equity market return data from 1836 to 2003, and argues the insignificant relation between risk and return documented in the previous literature that could be due to small sample problems. Using monthly data for the USA and annual data for the UK from 1871, Goetzmann and Jorion (1995) find only weak evidence of dividend yield predictability on long-horizon returns, whereas Goetzmann, Ibbotson, and Peng (2001) reach a similar conclusion using a new data set of the US market for the period of 1815–1925. In addition, they test for time-varying volatility using GARCH estimation and confirm earlier empirical evidence that positive stocks and negative stocks have different predictability for future volatility. Using a sample size of over 300 years should allow us to examine the robustness of calendar anomalies with a strong increase in the power of our tests.

Research into calendar anomalies, which we discuss more extensively below, is one of the oldest strands in the finance literature, starting with Wachtel's study in 1942 on the January effect, and followed by many other, now classic, studies, including Rozeff and Kinney (1976); French (1980); Gibbons and Hess (1981); Lakonishok and Levi (1982); Roll (1983); Keim (1983); Reinganum (1983); and Ariel (1987). Ever since 1942, old and new calendar anomalies [like the other January effect (Cooper, McConnell, and Ovtchinnikov, 2006) and seasonal effects in the cross-section of stock returns (Heston and Sadka, 2007)] keep practitioners and academics intrigued. Swinkels and Van Vliet (2010) try to disentangle the different calendar anomalies. Ogden (2003) relates equity return patterns to the seasonality of macroeconomic variables and a recent paper by Ogden and Fitzpatrick (2010) shows that many other anomalies, like the failure risk anomaly, earnings momentum, and the book-to-market anomaly, may also be seasonal. Many papers now assume that there are seasonal anomalies, like the January effect, and try to explain them. We feel that our article contributes to the literature, as it takes a step back and asks the question—using these new historical data—of whether or not these monthly seasonal anomalies exist and, if so, when they emerge. For instance, explanations for a January effect should allow for a valid explanation as to why the January effect changed from being a relatively negative month before 1830 to a positive month thereafter. Thus, understanding whether, and if so which, calendar anomalies persist helps our understanding of the working of financial markets and the behavior of investors.

2. A Short Literature Review on Monthly Calendar Anomalies

We summarize the main findings of seasonality studies in Table I. Given the data we use, we focus on the UK market. In Panel A, we report sample periods, data sources, weighting methods, and index types used in all the seasonality studies for the UK stock market. We also quote the key statistical findings at the market index level of each study to facilitate the comparison with our results. The last column reports the main reasons given by the studies for the observed seasonality. For the US market in Panel B, we report the studies that are either the first to document a particular calendar effect, or are the first investigation of a particular sample period. Panel C summarizes the empirical findings available for other countries. The positive (negative) sign indicates a significant positive (negative) effect.

Table I. Summary of empirical findings

Table 1 summarizes the empirical studies for the calendar month seasonals. PW, VW, and EW refer to price-, value-, and equally weighted index, respectively. The statistics are reported at percentage value, bold numbers in Panels A and B, and + (−) sign in Panel C denote statistically significant effects reported in the study. *Referred studies in Panel C*: (A) Brown *et al*. (1982); (B) Gultekin and Gultekin (1983); (C) Berges, McConnell, and Schlarbaum (1984); (D) Kato and Schallheim (1985); (E) Van den Bergh and Wessels (1985); (F) Tinic, Adeshi, and West (1987); (G) Ho (1990); (H) Bouman and Jacobsen (2002); (I) Fountas and Segredakis (2002); (J) Zarour (2007); (K) Hong and Yu (2009); (L) Darrat *et al*. (2011); (M) Lean (2011).

Notes: [a]Summer refers to the deviation of mean returns during summer months (July through September for Northern Hemisphere countries, January through March for Southern Hemisphere countries) from the rest of the year.

[b]Hal refers to the difference in mean returns between November through April and May through October.

[c]The result is the deviation of calendar month returns from the annual average modeled from a detrended index using first-order differencing.

[d]Mean returns for 12 calendar months, deviation of November–April returns from May to October returns for Hal.

[e]Statistics are not provided in the study.

[f]Value-weighted Cowles price index for the period 1910–25, equally weighted NYSE index for the period 1926–74.

[g]The two 6-month periods are October through March and April through September.

Empirical studies	Data period used	Country/ data source	Weighting	Index type	Statistic type	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Summer[a]	Hal[b]	Reasons Given for Seasonality
						Monthly seasonals														
Panel A: UK evidence																				
Gultekin and Gultekin (1983)	1959–79	Capital International Perspective Indices	VW	Price return	Mean	3.41	0.69	1.25	3.13	−1.21	−1.69	−1.11	1.88	−0.24	0.80	−0.61	2.06			Tax-loss selling
Reinganum and Shapiro (1987)	1956–65	London Share Price Data Base	EW	Total return	Mean	1.15	0.94	1.24	3.05	0.39	−0.79	−0.04	2.84	0.33	2.00	0.46	1.93			Tax-loss selling: seasonality is not detected in pretax period, but positive January and April effects are present posttax period
	1966–80					5.38	1.10	0.51	3.91	−0.21	−1.01	0.33	0.53	0.74	0.51	−1.01	1.32			
Corhay, Hawawini, and Michel (1987)	1969–83	London Stock Price Data Base	EW	Total return	Mean	5.49	2.21	0.73	4.19	−0.48	−1.39	1.22	1.13	−1.04	−0.07	−0.06	1.62			No January seasonal in risk premium and a positive risk premium only in April estimated based on CAPM
Clare, Psaradakis, and Thomas (1995)	1955–90	FTSE A All Share Index	VW	Price return	Deviation[c]	2.00	−0.33	−0.44	2.21	−1.34	−0.85	−0.90	0.76	−1.64	−1.34	−0.69	1.68			Seasonality is not related to size or seasonal variation

(continued)

Table I. Continued

Empirical studies	Data period used	Country/ data source	Weighting	Index type	Statistic type	Monthly seasonals												Summer[a]	Hal[b]	Reasons Given for Seasonality
						Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec			
																				of market risk measured using a GARCH(1,1)-M model.
Dimson and Marsh (2001)	1955–99	London Business School's Share Price Database	VW	Total return	Mean	2.83	0.64	1.05	2.67	−0.46	−0.71	−0.24	0.91	−0.65	−0.10	0.08	1.86			Seasonality appears in all size-sorted portfolios suggesting size-related explanations may not apply to the UK market
Choudhry (2001)	1870–13	NBER website	VW	Price return	Mean	1.13	−0.09	−0.69	0.08	−0.18	−0.25	−0.39	0.11	0.16	−0.20	−0.06	0.06			The January effect is not a small firm effect and not related to tax pre World War I period
Bouman and Jacobsen (2002)	1970–98	MSCI Reinvestment Indices	VW	Total return	Mean/ deviation[d]	3.10	0.80	−0.30	1.90	−1.40	−1.50	−0.10	0.00	−1.60	−0.80	1.30	1.20		2.02	Proxies for summer vacations are significantly related to the size of the effect. Other variables like interest rate, trading volume, size of agricultural sector are not related to the magnitude of Halloween effect among countries.
	1697–1969	Global Financial Data																	0.20	
Hong and Yu (2009)	1965–2005	Data stream	VW	Total return	Deviation													−1.50		Vacation induced lack of turnover and lower stock returns during summer months
Darrat et al. (2011)	1988–2010	MSCI Country Indices	VW	Total return	Deviation	−1.01	−0.21	0.19	1.60	−0.50	−1.88	0.78	−0.29	−1.81	0.45	0.50	2.17			NA

Table I. Continued

Empirical studies	Data period used	Country/data source	Weighting	Index type	Statistic type	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Summer[a]	Hal[b]	Reasons Given for Seasonality
Panel B: US evidence																				
Wachtel (1942)[e]	1928–40	DJIA	PW	Price return		+												+		Tax-loss selling for the January effect. Explanations not given for the positive summer returns
Rozeff and Kinney (1976)	1904–74	NYSE	EW[f]	Total return	Mean	3.48	0.26	−0.16	0.63	−0.37	0.18	1.90	1.46	−0.52	0.07	0.71	0.47			Risk, Information Hypothesis
Jones, Pearce, and Wilson (1987)	1871–1917	Cowles Industrial Index	VW	Price return	Mean	1.44	0.36	0.33	1.04	−0.55	−0.28	−0.52	1.12	0.35	−0.01	0.54	−0.35			Significant January effect is presented in both pretax and posttax period, indicating nontax factor to explain the January effect
	1918–38					2.61	0.89	0.33	−0.33	0.28	−0.61	3.22	2.39	1.05	−2.52	−0.13	−0.61			
Bouman and Jacobsen (2002)	1970–98	MSCI Reinvestment Indices	VW	Total return	Mean/ deviation[d]	1.20	0.00	0.00	0.10	−0.20	0.10	−0.20	−0.70	−1.30	−0.60	0.60	0.90		0.93	
Ogden (2003)	1947–2000	NYSE stocks	EW	Total return	Mean excess return	3.37	0.56	1.18	0.35	0.09	−0.36	0.79	0.19	−0.56	−0.90	1.29	1.51		1.12[g]	An annual cycle of economic activities and risk conditions
			VW			1.26	0.28	0.95	0.62	0.26	0.17	0.62	0.21	−0.68	0.09	1.41	1.72		0.79[g]	
Hong and Yu (2009)	1962–2005	Data stream	VW	Total return	Deviation													−1.10		
Darrat et al. (2011)	1988–2010	MSCI Country Indices	VW	Total return	Deviation	−0.35	−1.18	0.47	1.32	0.82	−1.22	0.32	−1.79	−0.94	0.27	0.86	1.42			NA
Panel C: International evidence																				
(L)		Argentina													−	−	+			
(J)		Abu Dhabi																	+	
(A), (B), (L)		Australia				+			+			+	+				+			
(H), (L)		Austria												−			+		+	
(J)		Bahrain																	+	
(B), (H), (L)		Belgium				+			+					−			+		+	
(H)		Brazil																+	+	
(B), (C), (F), (H), (K), (L)		Canada				+				+				−			+		+	
(I)		Chile				+														

(continued)

Table I. Continued

Empirical studies	Data period used	Country/data source	Index Weighting type	Statistic type	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Summer[a]	Hal[b]	Reasons Given for Seasonality
					Monthly seasonals														
(M)		China																+	
(I)		Colombia														+			
(B), (L)		Denmark			+								−			+			
(J)		Egypt																+	
(K), (L)		Finland								−							−		
(H), (K), (L)		France						+		−			−			+	−	+	
(B), (H), (K), (L)		Germany			+							−	−			+	−	+	
(H), (I), (L)		Greece			+			+			+			−				+	
(G), (K), (L)		Hong Kong			+							−					−		
(L)		Indonesia											−			+			
(M)		India																+	
(H), (L)		Ireland			+		+	+		−			−			+		+	
(H), (K), (L)		Italy						+					−			+	−	+	
(B), (D), (G), (H), (M)		Japan			+			+		+						+		+	
(L), (J)		Jordan			+	−						−						+	
(G), (L)		Korea			+	−													
(J)		Kuwait																+	
(G), (H), (I), (L), (M)		Malaysia			+		−					−				+		+	
(B), (E), (H), (K)		Netherlands			+			+					−			+	−	+	
(L)		New Zealand				−		+			+		−						
(K)		Nigeria															−		
(B), (K), (L)		Norway			+								−			+	−		
(J)		Oman																+	
(J)		Palestine																+	
(G), (H), (K), (L)		Philippines			+							−					−	+	
(L)		Portugal								−									
(K)		Russia															−		
(G), (H), (L), (M)		Singapore			+							−				+		+	
(B), (H), (K), (L)		Spain			+			+					−				−	+	
(B), (H), (L)		Sweden			+							−	−					+	
(B), (H), (K), (L)		Switzerland			+								−			+	−	+	
(G), (H)		Taiwan			+													+	
(K)		Thailand														+			
(I), (L)		Turkey			+		−		−			−				+			

2.1 CALENDAR MONTH SEASONALS

Wachtel (1942) uncovers the January effect in the US stock market as early as 1942. Interestingly, at least from our perspective, he documents it in a short sample from 1928 to 1940. However, as studies on seasonal behavior of stock market returns do not receive much academic attention at the time, it takes until 1976 before studies on the January effect become popular. In 1976, Rozeff and Kinney investigate the presence of seasonality in the USA. Their study made the January effect popular among academics using a relatively long sample of 70 years of NYSE index data from 1904 to 1974. Subsequently, many other studies document a January effect all over the world, albeit in generally relatively small samples, as shown in Table I.

To date, there is no conclusive evidence on what causes this January effect. In Wachtel's original study, he proposes five possible causes for the January effect: (i) tax-loss selling; (ii) unusual cash demand around Christmas; (iii) a pre-Christmas holiday effect; (iv) the anticipation of better business in Spring; and (v) a positive feeling about the coming new year. The tax-loss selling explanation[4] subsequently becomes the most widely investigated hypothesis, especially after Keim (1983) shows the January effect in the US market to be size related and concentrated in the small firms. The USA evidence generally supports the tax-loss selling hypothesis (see, for instance, Reinganum, 1983; Roll, 1983; Schultz, 1985; Jones, Lee, and Apenbrink, 1991; Poterba and Weisbenner, 2001; Starks, Yong and Zheng, 2006). At the same time, these studies cannot rule out the validity of other alternative explanations, like window dressing, the information hypothesis, the liquidity hypothesis, and optimistic expectations.[5]

[4] The tax-loss selling hypothesis states that downward pressure on stock prices might be induced at year end by investors selling the losing stocks with the intention of realizing capital losses against their taxable incomes. The abnormally high January return is the effect of the stock price rebounding to its equilibrium level when the selling pressure stops at the beginning of the year.

[5] The window dressing hypothesis is supported by Haugen and Lakonishok (1987), Lakonishok *et al.* (1991), and Ng and Wang (2004). It refers to the phenomenon when fund managers sell losing stocks prior to the disclosure of their portfolio holdings, typically at year end to impress investors, and buy the stocks back after the disclosure. The information hypothesis, discussed in Rozeff and Kinney (1976), Keim (1983), and Barry and Brown (1984), suggests that the January effect is caused by inappropriate modeling of risk: the market fails to account for the increased uncertainty in January due to the impending release of important information for the firms with a December fiscal year end. A related study, Kim (2006), constructs an earning information uncertainty risk factor that explains the January effect in the US market. The liquidity hypothesis proposed by Ogden (1990) argues that the January effect stems from the increased demand for stocks caused by liquid cash injection from year end salaries, bonuses, and dividend payments. The optimistic

Meanwhile, earlier seasonality studies outside the USA, primarily investigated as robustness checks for the tax-loss selling hypothesis and the January effect, suggest that the January effect is prevalent. However, these studies also find that tax-loss selling may only partially account for the January effect. In particular, Brown *et al.* (1982) find that Australian stocks during the period from 1958 to 1981 exhibit higher returns not only in July (in line with the tax-loss selling as the fiscal year ends in June), but also in December, January, and August. Using monthly data of value-weighted stock market indices of 17 industrialized countries from 1959 to 1979, Gultekin and Gultekin (1983) show the presence of the January effect in all 17 countries and an April effect for the UK market.[6] With the only exception of Australia, their finding is in support of the tax-loss selling hypothesis. Berges, McConnell, and Schlarbaum (1984) show, however, that the January effect in the Canadian stock market is present both before and after the introduction of capital gain tax in 1973 using 30 years data from the 1950s on. In addition, Tinic, Adeshi, and West (1987) find no seasonality in stocks traded by foreign investors and Canadians who were subjected to taxation before 1972, indicating that tax-loss selling cannot fully explain the January effect. In the Netherlands, Van den Bergh and Wessels (1985) find a January effect in the Dutch stock market for the period 1966–82 even though capital gains are not taxed. Although individual investors are not subject to capital gain taxes in Japan and the corporate fiscal year end varies among firms, Kato and Schallheim (1985) report both a January and a June effect for the Japanese stock market from 1952 to 1980. Their study inclines to support the alternative liquidity and information hypothesis.

For our study, the UK evidence is interesting. Reinganum and Shapiro (1987), using monthly data from 1955 to 1980, find support for the tax-loss selling hypothesis. They document both a January[7] and an April effect after the introduction of capital gain taxes in April 1965, whereas they detect no

expectation hypothesis suggested by Ciccone (2011) claims that the turn of the year is a time of renewed optimism that bids up the stock price in January. In addition, Anderson, Gerlach, and DiTraglia (2007) finds behaviorally related explanations are supported by laboratory tests.

[6] As the tax year ends on 5 April in the UK, an April effect is consistent with the tax-loss hypothesis.

[7] A January effect might be caused by international stock market integration; see Gultekin and Gultekin (1983) for evidence of the January effect in capital markets around the world. In addition, Reinganum and Shapiro (1987) suggest that the January effect in the UK stock market is driven by corporations that have a tax year ending at the end of December.

seasonality in the pretax period. In addition to the higher January and April returns, a later study by Clare, Psaradakis, and Thomas (1995) also reports high December returns and low September returns in the UK stock market during the period of 1955–90. With the benefit of cross-sectional data, studies show that the January effect in the UK (Dimson and Marsh, 2001) and Australia (Brown *et al.*, 1982) is a market-wide phenomenon, unlike in the USA, the anomaly in these countries is not related to firm size.

For emerging markets, Ho (1990) confirms the presence of the January effect in 7 out of 10 Asia Pacific markets. Fountas and Segredakis (2002) investigate monthly seasonality in 18 emerging markets and find a significant January effect in Chile, Greece, and Turkey, relatively high December returns in Colombia and Malaysia, and low October returns in Greece. A recent study (Darrat *et al.*, 2011) updates the monthly seasonalities in 34 equity markets, including the USA and the UK. Using a more recent sample period from 1988 to 2010, they find an absence of the January effect in all except three countries in the sample (Denmark, Ireland, and Jordan). Moreover, many stock market studies reveal significantly higher returns in April and December, whereas lower returns in June, August, and September.

2.2 HALLOWEEN EFFECT

The Halloween effect, or Sell-in-May effect, refers to the notion that stock market returns tend to be higher from November through April than from May through October. It originates from an old European market wisdom first investigated empirically by Bouman and Jacobsen (2002) using 37 countries' monthly return indices. They show that the Halloween effect is present in 36 stock markets, and statistically significant in 20 of those markets. Andrade, Chhaochharia, and Fuerst (2012) find that in an out-of-sample (1998–2012) period, all 37 of these countries in the original study have performed better in November through April than during the remainder of the year and 14 have done so significantly. In addition, Jacobsen, Mamun, and Visaltanachoti (2005) show that the Halloween effect is a market-wide phenomenon, which is not related to the common anomalies, such as size or Book to Market ratios and/or dividend yields. Jacobsen and Visaltanachoti (2009) investigate the Halloween effect among US stock market sectors and find substantial differences across sectors.

Zarour (2007) studies the Halloween effect in Arabic stock markets and Lean (2011) considers markets in Asia. Zarour (2007) finds that the Halloween effect is present in 7 of the 9 Arabic markets in the sample period from 1991 to 2004. Lean (2011) investigates six Asian countries for the period 1991–2008, and shows that the Halloween effect is only significant in Malaysia and Singapore, if modeled with OLS, but that three additional countries (China, India, and Japan) become statistically significant when modeled allowing for time-varying variance.

There are a number of explanations doing the rounds for what may cause this effect. Bouman and Jacobsen (2002) examine a large number of possible explanations. However, they can rule out many and their findings incline to support the vacation-induced change in risk aversion or liquidity hypothesis as a likely candidate. Interestingly, Hong and Yu (2009) report a similar seasonal trading pattern that turnovers are significantly lower over a 3-month period (July–September for Northern Hemisphere countries and January–March for Southern Hemisphere countries), which they attribute to investors taking summer vacations away from the stock market. In addition, they document significantly reduced summer returns in 15 out of 51 stock markets studied in their sample.

However, there are many other possible explanations. Ogden (2003) reports a similar seasonal pattern in the US stock returns. He finds that the mean excess return during October through March is significantly higher than the return from April through September and suggests an annual cycle view of economic activities and risk conditions. Gerlach (2007) attributes the significantly higher 3-month returns from October through December in the US market to higher macroeconomic news announcements during the period. Van der Gugten (2010) finds, however, that macroeconomic news announcements have no effect on the Halloween anomaly.

A number of studies also document a similar seasonal pattern in various stock markets, however, based on the alternative mood-related theories. For example, the Seasonal Affective Disorder (SAD) effect in Kamstra, Kramer, and Levi (2003), and the temperature effect in Cao and Wei (2005), are highly correlated with the Halloween effect, as shown by Jacobsen and Marquering (2008). However, as Jacobsen and Marquering (2008) point out correlation is not causation, therefore it is hard to distinguish between these explanations. Moreover, the validity of particularly the SAD paper by Kamstra, Kramer, and Levi (2003) has been strongly criticized by a number of studies. For instance, Kelly and Meschke (2010) show the model used in Kamstra, Kramer, and Levi (2003) is misspecified, due to a misreading of the evidence in the psychological literature regarding the timing of changes in mood. Kelly and Meschke (2010) then show that this misspecification drives the findings in Kamstra, Kramer, and Levi (2003).

3. Data

We obtain a 317-year index of monthly UK stock prices compiled by Global Financial Data from several different sources. Starting from 1693, the index basically covers the entire trading history of the UK equity market.[8] Table II summarizes the sources.

The index consists of stocks of the East India Company, the Bank of England, and the South Sea Company for the first 110 years. From a twenty-first century perspective, this may seem strange but in the eighteenth century, these three stocks essentially were the market.[9] These were the only stocks that traded on a daily, or at least weekly, basis before 1800. Other stocks could go an entire year without a price change.[10] Shea (2000) documents the total observable value of equity in the eighteenth century in relation to these three big companies. This confirms their relative importance at the time. Before 1810, the market share measured in market value of the three companies ranges between 98.50% at the beginning of the eighteenth century to 92.10% toward 1810. Mirowski (1981) examines surviving financial reports of some investing companies, indicating that their major investments were unanimously in these particular companies. He notes:

The relative insignificance of securities not linked to the government or the three big companies (Bank of England, East Indies and South Sea) in the eighteenth century is also supported by surviving evidence from much smaller balance sheets. The Scotch Mines Company's balance sheet shows that in 1773 the main assets were Bank of England securities (58 percent), East India Company annuities (31 percent), and bills of exchange (2 percent). No other company's shares were included.

Mirowski also constructs an annual index consisting of up to eight stocks[11] for the eighteenth century. If we compare this with our index, this allows us to evaluate to what extent the three big stocks were a good reflection of total market activity during this century. Figure 1 shows that, on an annual basis,

[8] Great Britain switched from the Julian calendar to the Gregorian calendar in September 1752. This change results in an omission of 11 days. Wednesday, September 2, 1752 was followed by Thursday September 14, 1752. Since our data are at monthly frequency, 11 days change within September should not have any effect on our results.

[9] Of course a three stock index might have a higher variance than more diversified indices of later periods and make estimates noisier; however, we show in our robustness tests that this hardly seems to affect our overall conclusions.

[10] Private correspondence with Bryan Taylor of Global Financial Data.

[11] Among the eight stocks, only the three companies included in our index have a continuous record for the whole of the eighteenth century.

Table II. Sources and descriptive statistics of subindices used to construct the Global Financial Data index

Dates	Source	No. of stocks	Companies/types	Weighting method	Mean (SD) (%)
1693	Thorold Roger, A history of prices in England (1693–97);	1	East Indies stock	–	−0.32 (5.94)
1694 to August 1711	Larry Neal, The rise of financial capitalism (1698–January 1811)	2	Bank of England and East Indies Stock	Equally weighted	
September 1711 to January 1811		3	Bank of England, East Indies Stock, and South Sea Stock	Equally weighted	0.03 (3.88)
02/1811 to 12/1850	Rostow's Total Index (Gayer, Rostow and Schwartz, 1975)	63	Canals, docks, waterworks, insurance, gas light and coke, mines, railways, and banks	Value weighted	−0.05 (4.19)
January 1851 to June 1867	Hayek's Index (Gayer, Rostow and Schwartz, 1975)	Unknown	Canals, docks, waterworks, gas light and coke, British mines, railways, and miscellaneous companies	Equally weighted	0.13 (1.94)
July 1867 to December 1906	London and Cambridge Economic Service Index	25–75	Broad based, but does not include bank, discount companies, insurance, and railways	Equally weighted	0.1 (1.52)
January 1907 to May 1933	Banker's Magazine	287	Broad based, virtually all stocks quoted on the exchange	Value weighted	−0.13 (2.51)
June 1933 to March 1962	Actuaries General Index	30 industrials	Blue Chip index represents several industries, including Financial Stocks, Commodities and Utilities, but excluded Debentures and Preferred Shares	Value weighted	0.4 (3.98)
April 1962 to December 2009	Financial Times-Actuaries All-Share Index	500 industrial companies	Broad based, represents 98–99% of capital value of all UK companies	Value weighted	0.58 (5.48)

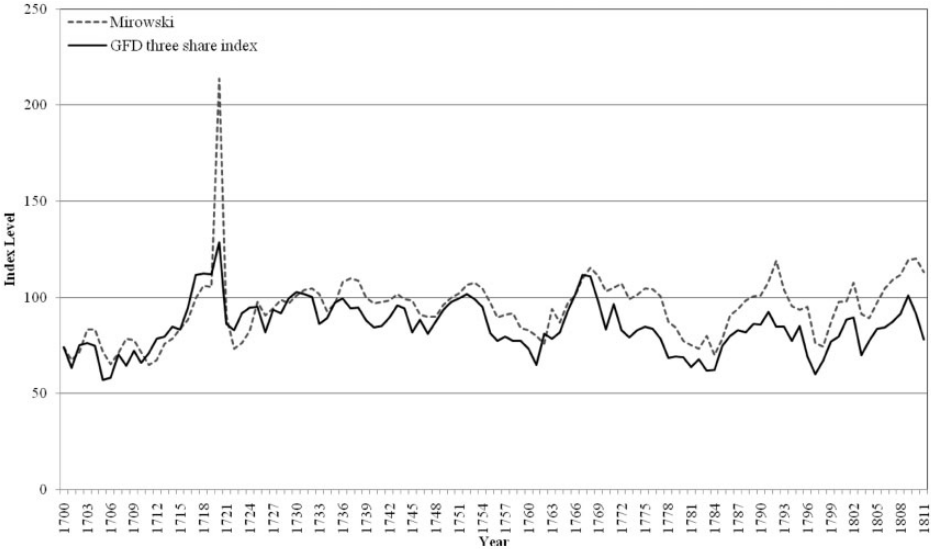

Figure 1. The Global Financial Data (GFD) three share index compared with Mirowski (1981) annual share price index (1700–1811).

the index based on these three big stocks seems well in line with the broader market index calculated by Mirowski (1981).

For the first half of the nineteenth century, the index adopts Rostow's total index (1811–50) and Hayek's index (1851–67), which are sourced from Gayer, Rostow, and Schwartz (1975). Both indices are broad based and favor large and frequently traded companies. The Rostow's total index represents one-third of the companies officially listed in the market. For the second half of the nineteenth century, the index uses the London and Cambridge Economic Service index constructed by Smith and G.F. Horne, which is the most widely studied index for the pre-World War I period. The Banker's Magazine index applies for the period from 1907 to 1933. It is the broadest index of London shares for the period. The stock market ceased trading for 5 months from August 1914–December 1914. We treat the data for this period as missing. The index consists of the Actuaries General Index from 1933 to 1962, and the Financial Times-Actuaries All-Share index, which covers ~98–99% of the capital value of all UK companies from April 1962 onward.

Some of these subindices are equally weighted, whereas others are value weighted. This might affect our estimation results, as the equally weighted indices will put relatively more weight in smaller companies. In our robustness section, we show that our results are not affected, if we replace all series by value-weighted indices wherever possible.

Most of these subindices are frequently used in other empirical studies; for example, Shiller (1989); Goetzmann (1993b); and Goetzmann and Jorion (1995). While the series does not include dividends, we show in our robustness tests that this does not seem to affect our overall results.[12]

4. Monthly Seasonality

Are stock returns in different months significantly different from each other? To study the potential effects of sample sizes on monthly stock returns, as discussed in Lakonishok and Smidt (1988), we first consider the full sample and also divide it into three (roughly) 100-year subperiods and six subperiods of around 50 years. This allows us to examine the monthly stock return seasonality with relatively large sample sizes, while still being able to detect any trends and persistent patterns over time. Table III reports the results for the general seasonality tests, as well as basic statistical characteristics of the returns for each calendar month. We also report basic characteristics for winter months (November through April) and summer months (May through October) defined by the Halloween effect and for the entire year over the various sample periods.

The latest 100- and 50-year subsamples enable us to confirm the findings of most earlier studies,[13] and the other two (and a half) centuries data can be safely treated as fresh data for out of sample tests over a longer time period, as they have not been studied before in relation to seasonal anomalies.

Overall, the average monthly return over the entire sample is only 0.12% (1.44% per year), which is relatively low, but this is due to the negative average returns during the first 150 years.[14] We observe an increasing

[12] Global Financial Data does not have a reliable long series including dividends before 1929. The only series available relies on the Bank of England stock mostly before 1922 and assumes a dividend yield for the next 7 years, however, even with that series the main conclusions in our article remain unaffected.

[13] Seasonality studies for the US market include earlier periods [i.e., the sample period in Wachtel (1942) starts from 1927, in Rozeff and Kinney (1976) from 1904, in Schultz (1985) from 1900, and in Jones, Pearce, and Wilson (1987) from 1871]. Sample periods in seasonality studies of the UK market focus on the latest 50-year subperiod of our sample. For example, Gultekin and Gultekin (1983) examine UK data from 1959 to 1979, Corhay, Hawawini, and Michel (1987) consider the period 1969–83, and Reinganum and Shapiro (1987) use the period 1955–80. A recent study by Dimson and Marsh (2001) investigates the period from 1955 to 1999, and Darrat *et al.* (2001) tests for the period 1988–2010.

[14] Negative capital gains in the long run may seem surprising nowadays, however, during the eighteenth and the nineteenth centuries dividends were relatively more important.

Table III. Seasonality tests and descriptive statistics of seasonal returns

Table III reports average return (percentage), standard deviation (percentage), skewness, and kurtosis for each calendar month, winter months (November through April), summer months (May through October), and entire year. The sample is subdivided into three subperiods of around 100-year intervals and six subperiods of 50-year intervals. Seasonality is tested using a Kruskal and Wallis (K–W test) rank-based non-parametric equality test and parametric joint significance test. The *F*-statistics reports the joint significance of the regression parameter α_2–α_{12} from the regression $R_t = \alpha_1 + \alpha_2 D_{2t} + \alpha_3 D_{3t} + \cdots + \alpha_{12} D_{12t} + \varepsilon_t$, where α_1 is the average return of January, and α_2–α_{12} represent the differences between January returns and the returns of the other months. ***denotes significance at the 1% level; **denotes significance at 5% level; *denotes significance at 10% level.

	January			February			March		
Sample period	Mean (SD)	Skew	Kurt.	Mean (SD)	Skew.	Kurt.	Mean (SD)	Skew.	Kurt.
1693–2009	0.69 (5.10)	4.90	51.28	0.09 (3.21)	0.46	9.49	−0.03 (3.73)	0.63	21.56
100-year Interval									
1693–1800	−0.60 (3.74)	−2.01	13.97	0.20 (3.04)	−0.48	8.91	0.11 (4.46)	2.11	24.34
1801–1900	1.34 (5.79)	7.53	67.79	−0.05 (2.47)	−2.07	11.08	−0.33 (2.14)	0.03	0.95
1901–2009	1.35 (5.37)	3.99	31.38	0.10 (3.92)	1.43	7.59	0.11 (4.07)	−1.50	9.60
50-year interval									
1693–1750	−0.48 (4.72)	−2.12	10.06	0.10 (3.71)	−0.52	7.04	−0.28 (5.72)	2.15	17.34
1751–1800	−0.73 (2.18)	1.59	6.50	0.32 (2.05)	0.44	2.71	0.56 (2.24)	−1.15	4.14
1801–50	1.55 (7.98)	5.83	38.39	−0.25 (3.03)	−2.29	9.13	−0.50 (2.53)	−0.29	−0.66
1851–1900	1.12 (2.00)	1.75	5.52	0.14 (1.75)	0.54	0.63	−0.16 (1.68)	1.52	5.44
1901–50	0.86 (1.35)	−0.23	0.54	−0.50 (2.32)	−1.59	5.82	−0.49 (2.50)	0.60	2.79
1951–2009	1.75 (7.19)	2.97	16.98	0.60 (4.85)	1.32	4.59	0.62 (5.00)	−1.83	8.15

	April			May			June		
Sample period	Mean (SD)	Skew	Kurt.	Mean (SD)	Skew.	Kurt.	Mean (SD)	Skew.	Kurt.
1693–2009	0.49 (3.39)	−0.25	6.33	0.02 (4.11)	3.03	42.74	−0.12 (3.78)	3.44	42.94
100-year Interval									
1693–1800	0.31 (3.01)	−0.33	6.66	0.48 (5.41)	4.24	40.62	0.31 (4.58)	6.35	56.17
1801–1900	−0.40 (2.65)	−3.44	22.87	−0.22 (2.59)	−1.85	6.50	0.20 (2.05)	0.28	3.52
1901–2009	1.50 (4.05)	0.16	2.02	−0.21 (3.74)	−0.71	1.26	−0.85 (4.05)	−0.74	1.85
50-year Interval									
1693–1750	0.61 (3.70)	−0.57	5.17	1.09 (7.11)	3.35	24.33	0.61 (6.00)	5.21	35.18
1751–1800	−0.04 (1.92)	0.30	1.78	−0.23 (2.02)	−1.58	3.95	−0.04 (1.90)	1.21	4.85
1801–50	−0.60 (3.48)	−3.00	14.77	−0.25 (3.30)	−1.63	3.99	0.47 (2.36)	0.78	1.54
1851–1900	−0.21 (1.42)	0.10	0.74	−0.19 (1.64)	−1.61	6.92	−0.07 (1.65)	−1.72	7.05
1901–50	0.11 (2.79)	−0.98	2.93	0.12 (2.76)	−1.01	3.13	−0.94 (3.68)	−1.46	5.21
1951–2009	2.67 (4.57)	−0.09	1.49	−0.49 (4.40)	−0.50	0.33	−0.77 (4.37)	−0.40	0.41

	July			August			September		
Sample period	Mean (SD)	Skew	Kurt.	Mean (SD)	Skew.	Kurt.	Mean (SD)	Skew.	Kurt.
1693–2009	−0.31 (3.31)	−1.10	8.81	0.44 (3.25)	−0.09	2.78	−0.49 (5.62)	−7.07	91.72
100-year interval									
1693–1800	−0.45 (3.04)	0.13	2.97	0.73 (2.77)	0.46	1.98	−0.93 (8.15)	−6.52	60.47
1801–1900	−0.49 (1.90)	−0.24	0.85	−0.32 (1.94)	−0.49	2.27	−0.27 (2.19)	−1.20	5.15
1901–2009	0.00 (4.41)	−1.55	7.27	0.86 (4.36)	−0.46	1.34	−0.26 (4.68)	−1.36	2.71
50-year interval									
1693–1750	−0.34 (3.71)	0.02	1.93	0.71 (3.05)	0.26	1.75	−1.81 (10.95)	−4.89	33.67
1751–1800	−0.57 (2.03)	0.39	0.29	0.74 (2.42)	0.93	2.24	0.08 (1.95)	−0.81	1.56
1801–50	−0.94 (2.11)	−0.13	0.19	−0.78 (2.48)	−0.02	0.77	−0.86 (2.64)	−0.95	3.66
1851–1900	−0.05 (1.55)	0.17	1.94	0.14 (1.02)	0.27	−0.62	0.32 (1.44)	0.13	−0.20
1901–50	−0.18 (4.55)	−2.91	15.18	0.44 (3.25)	−0.14	2.45	0.40 (2.51)	−1.47	4.11
1951–2009	0.16 (4.32)	−0.26	−0.49	1.21 (5.11)	−0.63	0.76	−0.82 (5.90)	−0.95	0.71

(continued)

Table III. Continued

Sample period	October Mean (SD)	Skew	Kurt.	November Mean (SD)	Skew.	Kurt.	December Mean (SD)	Skew.	Kurt.
1693–2009	−0.50 (4.37)	−2.55	19.22	0.35 (3.86)	0.25	9.35	0.81 (3.22)	1.53	10.91
100-year interval									
1693–1800	−1.38 (4.99)	−2.75	18.56	0.17 (3.48)	−0.49	6.76	0.61 (2.51)	0.55	2.02
1801–1900	−0.12 (2.37)	0.74	6.01	0.36 (3.70)	3.01	24.89	1.00 (3.56)	2.76	19.33
1901–2009	0.02 (4.99)	−2.26	13.49	0.51 (4.35)	−0.99	2.83	0.82 (3.53)	0.58	3.24
50-year interval									
1693–1750	−1.95 (6.56)	−2.09	10.24	0.45 (3.78)	−0.37	7.23	0.80 (2.97)	0.48	1.08
1751–1800	−0.73 (1.87)	1.30	4.81	−0.16 (3.10)	−0.93	5.78	0.39 (1.83)	0.13	2.39
1801–50	−0.28 (2.81)	0.74	5.39	0.55 (4.88)	2.54	16.18	1.67 (4.55)	2.23	12.97
1851–1900	0.04 (1.83)	0.99	4.07	0.17 (1.96)	1.12	3.97	0.33 (2.01)	1.58	8.26
1901–50	−0.03 (3.08)	0.47	4.32	0.82 (3.27)	0.47	2.35	−0.43 (2.43)	−0.88	3.29
1951–2009	0.06 (6.19)	−2.30	10.17	0.24 (5.11)	−1.16	1.84	1.89 (3.97)	0.43	2.43

Sample period	Winter Mean (SD)	Skew	Kurt.	Summer Mean (SD)	Skew.	Kurt.	Annual Mean (SD)	Skew.	Kurt.
1693–2009	0.40 (3.81)	2.27	34.85	−0.16 (4.16)	−2.62	67.89	0.12 (4.00)	−0.51	54.70
100-year interval									
1693–1800	0.13 (3.44)	0.14	17.42	−0.21 (5.18)	−3.18	79.30	−0.04 (4.40)	−2.60	79.95
1801–1900	0.32 (3.64)	5.87	84.12	−0.20 (2.19)	−0.57	5.06	0.06 (3.01)	5.18	92.37
1901–2009	0.73 (4.27)	1.21	16.41	−0.07 (4.40)	−1.28	5.78	0.33 (4.36)	−0.10	11.02
50-year interval									
1693–1750	0.20 (4.19)	0.15	13.80	−0.28 (6.80)	−2.60	48.80	−0.04 (5.65)	−2.29	53.78
1751–1800	0.06 (2.28)	−0.29	4.88	−0.13 (2.08)	0.36	2.57	−0.03 (2.18)	0.00	3.89
1801–50	0.40 (4.81)	5.01	54.48	−0.44 (2.67)	−0.39	3.35	−0.02 (3.91)	4.76	65.01
1851–1900	0.23 (1.86)	1.26	4.78	0.03 (1.54)	−0.36	4.61	0.13 (1.71)	0.71	5.13
1901–50	0.06 (2.56)	−0.25	3.60	−0.03 (3.38)	−1.66	10.51	0.02 (2.99)	−1.27	9.63
1951–2009	1.30 (5.25)	0.99	12.06	−0.11 (5.12)	−1.09	3.77	0.59 (5.23)	0.00	8.31

Sample period	Seasonality Test K–W		F-Stat	
1693–2009	55.07	***	3.84	***
100-year interval				
1693–1800	59.75	***	3.06	***
1801–1900	41.21	***	1.84	**
1901–2009	35.20	***	2.76	***
50-year interval				
1693–1750	36.06	***	1.42	
1751–1800	50.59	***	3.87	***
1801–1850	36.90	***	2.34	***
1851–1900	23.81	***	1.93	**
1901–50	31.31	***	2.88	***
1951–2009	30.09	***	2.90	***

trend in average price returns over time, with the latest 50 years showing the highest average return. While the standard deviations of different sample periods do not have a clear pattern, the market in the nineteenth century

We see relatively high dividend payments (around 5% annually) in the first two centuries of our sample: the series, including dividends (not reported in the table) has monthly returns of 0.53% and 0.40% in the first two centuries, respectively.

seems to be less volatile than it does in the eighteenth and the twentieth centuries.

The last two columns report the results of the calendar month seasonality tests. We use both parametric and nonparametric tests. The latter is the Kruskal and Wallis rank-based test of equality. The null hypothesis is that all of the calendar months have the same continuous distribution and that the test statistic is approximately distributed as a χ^2 with 11 degrees of freedom. The alternative hypothesis is that at least 1 month has a different distribution. The parametric test examines the joint significance of parameters α_2–α_{12} from the following regression Equation (1):

$$r_t = \alpha_1 + \alpha_2 D_{2t} + \alpha_3 D_{3t} + \cdots + \alpha_{12} D_{12t} + \varepsilon_t \tag{1}$$

where r_t is the monthly continuously compounded index returns, and $D_{2t} \ldots D_{12t}$ denote dummy variables for February to December. The constant parameter α_1 is the average return for January, and the coefficient estimates α_2–α_{12} represent the differences between January returns and the returns in other months. If returns for each month of the year are the same, the parameters α_2–α_{12} should be jointly insignificant. Both tests reveal strong calendar month seasonality over all of our examined sample periods.

While our tests statistics indicate significant differences between months, these tests do not clarify which month contributes to this seasonality and whether it is the same month in different samples. Based on the literature, we expect to see higher returns in January, April, and December, whereas lower returns in September (Reinganum and Shapiro, 1987; Clare, Psaradakis, and Thomas 1995; Dimson and Marsh, 2001). Note that our results confirm these findings. For the subsample period 1951–2009, April, December, and January have the highest returns, whereas the average September return is the lowest during the period. The interesting question is whether we will find similar results in earlier subperiods.

Our evidence in Table III suggests that these patterns do not persist over time. January returns are negative and lower than for the other months in the first 100 years, with the best month over the 300 years being December rather than January. The overall performance for October seems similar to September (–0.50% versus –0.49% return per month), but the average October return is higher than September in the most recent 50 years. In Table IV, we test the statistical significance of the individual months in more detail, using the standard random walk regression with a dummy variable:

$$r_t = \alpha + \beta_m D_{\mathrm{mt}} + \varepsilon_t \tag{2}$$

Table IV. Calendar month effects and the Halloween effect: OLS regressions

Table IV presents the coefficients estimates (percentage) and the t-statistics of the regression in a form of $r_t = \alpha + \beta_m D_{mt} + \varepsilon_t$, where ε_t is the continuously compounded monthly returns, r_t is the dummy variable of the calendar month m (or the Halloween dummy that equals 1 if the month falls on the period November through April and 0 otherwise for the Halloween effect regression), α is the constant and D_{mt} is the error term. T-statistics are calculated based on Newey and West (1987) standard errors. The sample is subdivided into three subperiods of approximately 100-year intervals and six subperiods of 50-year intervals. ***denotes significance at the 1% level; **denotes significance at 5% level; *denotes significance at 10% level.

	January		February		March		April		May		June	
Sample period	β	*t-value*	β	*t-value*	β	*t-value*	β	*t-value*	β	*t-value*	β	*t-value*
1693–2009	0.62	2.14**	−0.04	−0.19	−0.16	−0.76	0.41	2.03**	−0.11	−0.46	−0.26	−1.18
100-year interval												
1693–1800	−0.61	−1.82*	0.26	0.90	0.16	0.55	0.38	1.02	0.56	0.98	0.38	0.84
1801–1900	1.40	2.52**	−0.12	−0.40	−0.42	−1.67*	−0.50	−1.77*	−0.31	−1.16	0.16	0.74
1901–2009	1.11	2.09	−0.25	−0.65	−0.24	−0.57	1.28	2.72***	−0.59	−1.87*	−1.28	−3.27***
50-year interval												
1693–1750	−0.48	−0.83	0.15	0.31	−0.26	−0.57	0.71	1.10	1.23	1.23	0.71	0.91
1751–1800	−0.76	−2.72***	0.39	1.43	0.65	2.36**	−0.01	−0.04	−0.21	−0.80	−0.01	−0.02
1801–50	1.72	1.60	−0.25	−0.46	−0.53	−1.30	−0.63	−1.21	−0.26	−0.56	0.53	1.54
1851–1900	1.08	3.51***	0.01	0.06	−0.32	−1.03	−0.37	−1.70*	−0.36	−1.39	−0.22	−0.99
1901–50	0.93	4.08***	−0.56	−1.48	−0.55	−1.75*	0.10	0.20	0.11	0.32	−1.04	−2.81***
1951–2009	1.26	1.34	0.01	0.01	0.03	0.04	2.27	3.98***	−1.18	−2.43**	−1.49	−2.29**

	July		August		September		October		November		December		Halloween	
Sample period	β	*t-value*	β	*t-value*	β	*t-value*	β	*t-value*	β	*t-value*	β	*t-value*	β	*t-value*
1693–2009	−0.46	−2.35**	0.35	1.85*	−0.67	−2.08**	−0.68	−2.74***	0.25	1.11	0.75	3.99***	0.56	4.04***
100-year interval														
1693–1800	−0.45	−1.56	0.83	2.73***	−0.98	−1.06	−1.47	−3.67***	0.23	0.79	0.71	2.43**	0.34	1.37
1801–1900	−0.60	−3.00***	−0.42	−1.92	−0.35	−1.33	−0.19	−0.77	0.33	1.13	1.03	3.14***	0.52	2.98***
1901–2009	−0.36	−0.94	0.58	1.47	−0.64	−1.53	−0.34	−0.81	0.19	0.44	0.54	1.34	0.80	3.37***
50-year interval														
1693–1750	−0.33	−0.72	0.82	1.66*	−1.93	−1.17	−2.08	−3.14***	0.53	1.28	0.92	1.86*	0.48	1.09
1751–1800	−0.59	−1.81*	0.84	2.61***	0.12	0.42	−0.76	−2.78***	−0.13	−0.37	0.46	1.91*	0.18	1.03
1801–50	−1.00	−3.31***	−0.83	−2.38**	−0.92	−2.28**	−0.29	−0.68	0.62	1.14	1.84	3.40***	0.84	2.68***
1851–1900	−0.20	−1.03	0.00	0.03	0.21	0.89	−0.10	−0.38	0.04	0.22	0.21	0.87	0.20	1.41
1901–50	−0.22	−0.38	0.47	1.16	0.42	1.48	−0.05	−0.15	0.88	1.41	−0.49	−1.20	0.10	0.39
1951–2009	−0.47	−0.96	0.68	1.09	−1.54	−2.27**	−0.59	−0.81	−0.39	−0.62	1.41	2.97***	1.40	3.75***

where r_t is the continuously compounded monthly index return, D_{mt} is the dummy variable for a particular month (or a Halloween dummy that equals 1, if month t falls in the period from November through April and 0 otherwise), α is the constant and ε_t is the error term. β_m shows the magnitude of the difference between the mean return of the month(s) of interest and the mean return during the rest of the year.

Table IV contains our coefficient estimates and t-statistics based on Newey–West standard errors for each calendar month and the Halloween effect. As before, we consider the full sample results and the 100- and 50-year subsamples. To ensure that our results do not depend on the choice of the specific 50-year subsamples and to detect possible structural breaks,[15] we plot 50-year rolling window estimates for each of the 12 calendar month effects and the Halloween effect with their corresponding confidence bounds over the full sample in Figure 2.[16] These plots also illustrate how these monthly patterns vary over time and how—as Lakonishok and Smidt (1988) point out—the relatively large standard errors even with a sample size as long as 50 years make it difficult to infer statistical significance.

Over the entire sample period (Table IV), the December, January, April, and August returns are significantly higher than the returns for the rest of the year. Despite this, however, none of these months persistently outperform the market. December comes close with negative coefficient estimates only in the subperiod 1901–50. Even the well-known January effect appears only in the second half of our sample. Intriguingly, on average, the January returns are significantly lower, rather than higher, during the eighteenth century. Before 1850, a strong positive December effect dominates the market, which disappears as the January effect emerges in the nineteenth century.

Figure 2 shows this shift in January returns more clearly. January returns are rarely higher than the average months until 1820–30. (Note that the extremely high January returns exhibited in the 1820s are partially caused by the Panic of 1825[17] leading to an upward shift and, subsequently, to a

[15] We also performed formal structural break tests. While these confirm the results we find, they tend, however, to be sensitive to data trimming assumptions and, more importantly, did not provide the insight and detail these rolling regressions provide.

[16] We use 50 years to reduce the effect of outliers and to make our results comparable with the GARCH estimates we use in our robustness tests. Jacobsen and Dannenburg (2003) show that for reliable GARCH estimates in monthly data, one needs around 50 years of monthly observations.

[17] During the period, our index shows that the price level started to rise dramatically, by >20% per month, from November 1824, and had the largest increase of 54% in January 1825. Price levels remained high for 3 months and then sharply dropped back to the original

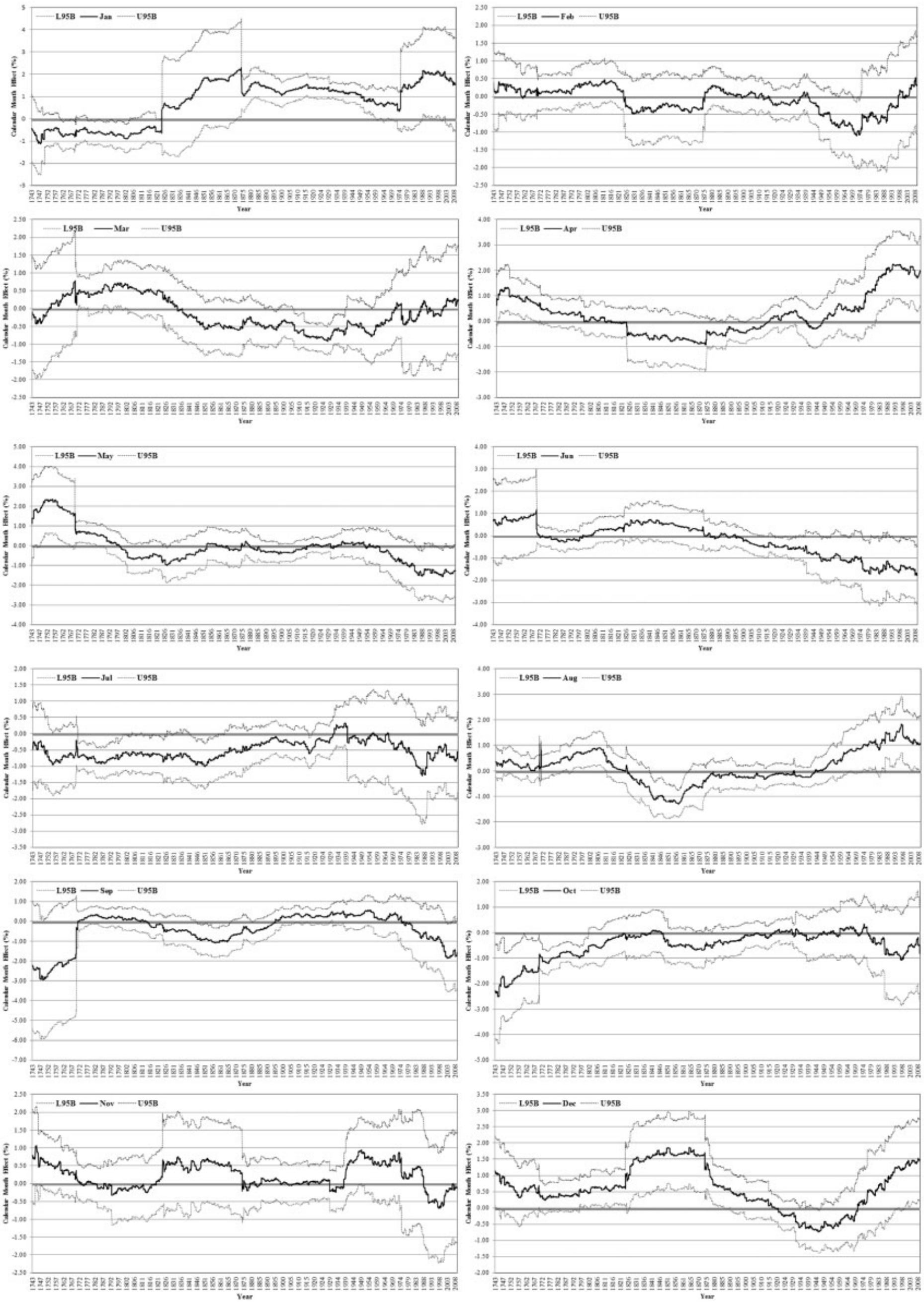

Figure 2. A 50-year rolling window OLS regressions of estimates for the 12 calendar month effects and the Halloween effect. The dark solid line indicates the coefficient estimates of the effect, the light dotted lines indicate the upper and the lower 95% bounds calculated based on Newey and West (1987) standard errors.

Figure 2. Continued.

strong downward shift in the rolling regression estimates. In our robustness tests, we perform an outlier robust regression and find that these outliers do not tend to influence our overall findings.) Only around 1830, do January returns become higher than those of other months and these higher returns continue to the end of the twentieth century. The higher January returns start from the mid-1830s, if we exclude the extreme price behavior in 1825. It is, however, not clear what causes this January effect, as a tax-loss selling explanation does not seem feasible. In particular, the UK capital gains tax was not imposed until 1965 with a tax year end of April, and income tax was first introduced in 1799, but repealed in 1816 and not reintroduced until 1842, however, neither of these periods coincide with the emergence of the January effect in the 1830s. Thus, tax-loss selling by individual investors with an April tax year end, or corporations and traders with a December tax year end, cannot explain the effect. In addition, income tax was not prevalent in other countries during the nineteenth century. For example, the USA introduced the War Revenue Act in 1917. Therefore, the emerging January effect cannot have been carried over from the USA. Tax-loss selling by foreign traders is also unable to explain the emergence of the January effect in the 1830s. An alternative explanation would be that the January effect is imported from the US market for a different reason; however, January returns in the USA are significantly below average up

level within a year. This price behavior is consistent with the description in Glasner (1997, p. 511), "...a speculative fever which seems to have begun in late 1824. They included a widespread feeling of optimism at the time, a general shortage of investment vehicles resulting from the decrease in the interests on bonds, an excess demand for several commodities, and the opening up of investment opportunities in South America.... At the beginning of 1824, there were 154 joint stock companies with capital of £48 million. An additional 624 such companies were either started or proposed during the next two years, 127 of which survived the crisis and were still in operation in 1927. The crash in the real sector followed that of the financial sector, with the bottom being reached in 1826."

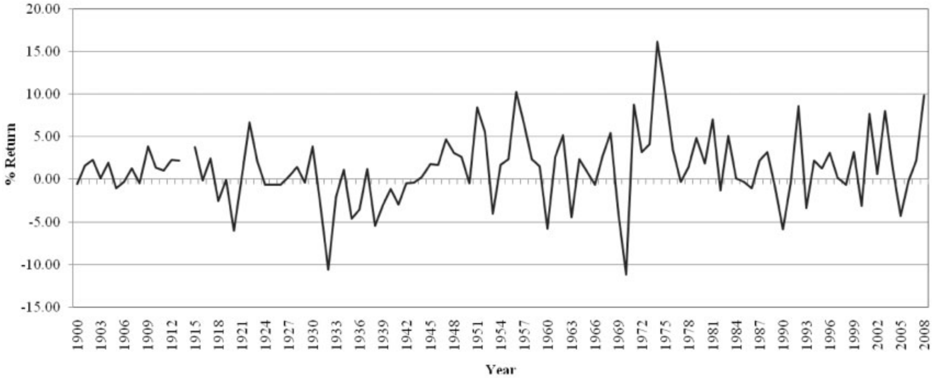

Figure 3. Stock market return difference between April and the average of the 11 other months (GFD index).

to 1870 and change thereafter.[18] The emerging January effect around this time in both the UK and the USA might offer some support for the Christmas hypothesis introduced by Wachtel (1942) as an explanation for the January effect, as the UK started officially celebrating Christmas in 1835 or 1837,[19] and in the USA Christmas was declared a legal holiday in 1870 by President Ulysses S. Grant. Clearly, the evidence we report here is speculative, but suggests that the Christmas hypothesis put forward by Wachtel in 1942 may deserve more attention.

In the UK, a capital gains tax was introduced on April 6, 1965. The results of, for instance, Reinganum and Shapiro (1987) suggest that this leads to the introduction of higher April returns from that point on. They find no

[18] Estimations based on extended S&P 500 composite price index data obtained from Global Financial Data over the period 1791–2009. The results are not reported here but available on request from the authors.

[19] Christmas becomes a national holiday in 1835 according to the website http://www.johnowensmith.co.uk/histdate/, but other sources (http://www.historic-uk.com/HistoryUK/England-History/VictorianChristmas.htm) suggest that the Christmas holiday is introduced later, in 1837: "Before Victoria's reign started in 1837 nobody in Britain had heard of Santa Claus or Christmas Crackers. No Christmas cards were sent and most people did not have holidays from work. The wealth and technologies generated by the industrial revolution of the Victorian era changed the face of Christmas forever . . . the wealth generated by the new factories and industries of the Victorian age allowed middle class families in England and Wales to take time off work and celebrate over two days, Christmas Day and Boxing Day. Boxing Day, December 26th, earned its name as the day servants and working people opened the boxes in which they had collected gifts of money from the '"rich folk'". Those new fangled inventions, the railways allowed the country folk who had moved into the towns and cities in search of work to return home for a family Christmas."

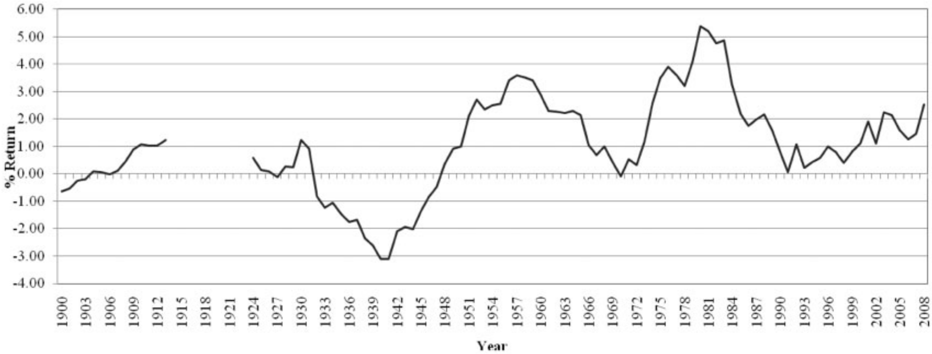

Figure 4. GFD stock market return difference between April and average of the other 11 months (10-year moving average).

seasonality in monthly UK returns in the 10 years prior to the introduction of capital gains tax. Having the benefit of a longer sample, we can revisit their evidence. The plot based on a rolling window of 50 years in Figure 2 suggests that around this time average April returns indeed do become higher. The evidence is, however, less conclusive if we plot annual April returns minus the average returns of the other 11 months (Figure 3), and a 10-year moving average of April returns minus the average returns of the other months (Figure 4) for the period 1900–2009 when April effect becomes positive.

Positive April returns occur frequently; however, it is not definite that the outperformance occurs only in the period after the imposition of the capital gains tax in 1965. In fact, the smoothed graph using a 10-year moving average reveals that the rising trend starts from the 1940s onward. This suggests that it may not necessarily be the capital gains tax that causes these higher April returns to emerge.

Table III shows that the average returns for October, September, and July are frequently negative. Table IV reveals that the relatively worst months are October and July, which significantly underperform the other calendar months over the whole sample period. They also persistently underperform in all subperiods. Although the results are not statistically significant for all subsamples, the coefficient estimates are unanimously negative. The average return for October over the whole sample period is 0.68% lower than the other months' averages. For July, this is 0.46%. However, the statistical significance weakens after the 1850s. The plots confirm that this is not a result of the specific sample periods we used. Based on a 50-year window, one rarely sees positive estimates for both July and October; however, for September things are different.

Our results confirm the low September returns reported by Clare, Psaradakis, and Thomas (1995) for the period of 1955–90. With the benefit of a longer sample period, however, we are able to show that the pattern is not persistent and that the September mean returns are actually higher than the returns during the other months for 3 out of the 6 50-year subperiods, although the difference is not statistically significant. Also the September plot in Figure 2 shows that over 300, it is hard to conclude that stock returns show a negative September effect.

The Halloween effect seems relatively robust over time. Six monthly winter returns tend to be on average 3.4% higher than 6 monthly summer returns measured over 300 years. In the first half of the twentieth century, this drops to around half a percent to later increase to 8.4% in the last 60 years of our sample. However, there are long periods when the effect does not show up significantly. And the point estimates even indicate a reversed effect in the early twentieth century, although not significantly so.

All-in-all, our evidence suggests that findings regarding many monthly anomalies may be less robust and very time dependent. This might either mean that there are no monthly seasonal effects, or alternatively that these monthly seasonals are themselves time varying. Unfortunately, in the latter case, the evidence from the past 300 years suggests that the monthly seasonals are varying over time with a speed that we might never be able to estimate whether they are real, or not, at least not with current estimation methods. If we require that coefficients need to be persistently negative, or positive, almost all of the time and be significant over the full sample, the only exceptions may be the negative July and October effects, and the Halloween effect. It is hard to find 50-year periods when these effects change signs. But based on the 50-year samples, it is difficult to conclude these effects are significantly present. We now check the robustness of these results against alternative specifications.

5. Robustness Checks

5.1 VOLATILITY CLUSTERING AND POSSIBLE OUTLIERS

In the first part of our sample, we use an index of only three stocks. This may increase volatility and reduce the power of our test statistics. Moreover, as monthly stock returns may also exhibit volatility clustering when we use Newey–West standard errors that are heteroscedasticity and autocorrelation consistent, this may reduce the power of our tests. To verify the impact the volatility might have, we first plot an annualized 5-year moving average standard deviation (Figure 5).

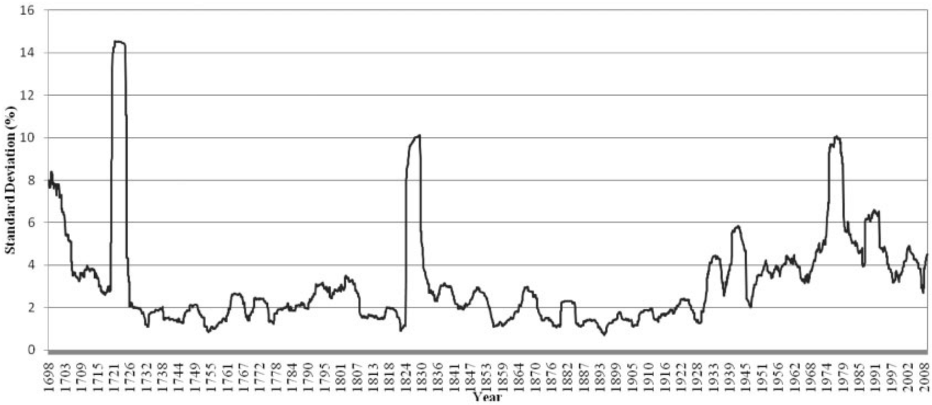

Figure 5. A 5-year moving standard deviation for the GFD stock market index.

This shows a couple of things. Indeed volatility is higher at the start of the sample, although it decreases to a low level even in the case of the three stocks. It also tends to spike, for instance, during the South Sea Bubble. This also seems to show up as wider confidence bounds in the rolling window regressions. Interestingly, volatility also seems to have increased on average in the twentieth century. Overall, because volatility is time varying and spikes occasionally, it may be good to verify robustness of our results controlling for both conditional heteroscedasticity and outliers using GARCH models and OLS robust regressions. Of course, the price we pay is that we have to impose a specific structure on the conditional heteroscedasticity and may accidentally exclude observations that were not outliers. As a robustness check, however, it may be good to make these assumptions.

For the GARCH model, we use a GARCH(1,1) model, as this simple parsimonious representation generally captures volatility clustering well in monthly data if we use a window of around 50 years or more (see, for instance, Jacobsen and Dannenburg, 2003). We estimate both assuming a normal distribution and *t*-distributed standard errors, but as the results are similar, we only report the former. We use the same mean equation with a dummy for the different months as used in our main regressions and reestimate the seasonal effects from Equation (3).

$$\begin{aligned} r_t &= \mu + \beta_m D_{\mathrm{mt}} + \varepsilon_t, \\ \varepsilon_t|\Phi_{t-1} &\sim N\left(0, \sigma_t^2\right), \\ \sigma_t^2 &= \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \sigma_{t-1}^2 \end{aligned} \tag{3}$$

Table V. Calendar month effects and the Halloween effect: GARCH(1,1) models

Table V presents the coefficients estimates (percentage) and the *t*-statistics of the calendar month effect and the Halloween effect estimated with GARCH(1,1) model: $r_t = \mu + \beta_m D_{\mathrm{mt}} + \varepsilon_t, \varepsilon_t|\Phi_{t-1} \sim N(0, \sigma_t^2), \sigma_t^2 = \alpha_0 + \alpha_1\varepsilon_{t-1}^2 + \alpha_2\sigma_{t-1}^2$, where r_t is the continuously compounded monthly returns, D_{mt} is the dummy variable for calendar month m (or the Halloween dummy that equals 1, if the month falls on the period November through April and 0 otherwise). The sample is subdivided into three subperiods of approximately 100-year intervals and six subperiods of 50-year intervals. ***denotes significance at the 1% level; **denotes significance at 5% level; * denotes significance at 10% level.

	January		February		March		April		May		June	
Sample period	β	*t-value*	β	*t-value*	β	*t-value*	β	*t-value*	β	*t-value*	β	*t-value*
1693–2009	0.23	1.74*	−0.05	−0.39	−0.05	−0.38	0.11	0.89	0.02	0.15	−0.28	−2.09**
100-year interval												
1693–1800	−0.74	−3.33***	0.08	0.33	0.54	2.65***	0.26	1.35	0.37	1.97**	−0.18	−0.71
1801–1900	0.71	3.24***	−0.02	−0.11	−0.22	−1.12	−0.38	−1.74*	−0.18	−0.84	−0.07	−0.30
1901–2009	0.85	2.37**	−0.30	−1.06	−0.53	−2.33**	0.76	2.99***	−0.17	−0.67	−0.85	−2.94***
50-year interval												
1693–1750	−0.60	−1.63	−0.28	−0.82	0.28	0.91	0.53	2.04**	1.29	3.79***	−0.27	−0.67
1751–1800	−0.82	−2.68***	0.33	0.95	0.67	2.04**	0.00	0.01	−0.21	−0.79	−0.14	−0.41
1801–50	0.54	1.23	0.25	0.44	−0.09	−0.21	−0.22	−0.45	−0.06	−0.12	0.34	0.70
1851–1900	0.82	3.64***	−0.13	−0.62	−0.31	−1.47	−0.41	−1.93*	−0.19	−0.85	−0.13	−0.55
1901–50	1.04	2.19**	−0.24	−0.78	−0.84	−3.23***	0.63	2.24**	0.31	1.08	−0.70	−2.24**
1951–2009	0.52	0.81	−0.48	−0.75	0.39	0.61	1.42	2.31**	−1.52	−2.91***	−1.31	−2.08**

	July		August		September		October		November		December		Halloween	
Sample period	β	*t-value*	β	*t-value*	β	*t-value*	β	*t-value*	β	*t-value*	β	*t-value*	β	*t-value*
1693–2009	−0.56	−4.35***	0.01	0.09	−0.01	−0.05	−0.24	−1.88*	0.28	2.39**	0.45	3.31***	0.32	4.44***
100-year interval														
1693–1800	−0.83	−4.00***	0.55	2.91***	0.08	0.37	−0.86	−3.77***	0.24	1.21	0.51	2.23**	0.26	2.25**
1801–1900	−0.39	−1.56	−0.61	−3.34***	−0.06	−0.28	0.12	0.61	0.39	1.94*	0.53	2.64***	0.38	3.19***
1901–2009	−0.45	−1.89*	0.41	1.65*	−0.07	−0.23	0.11	0.45	0.10	0.42	0.18	0.53	0.31	1.96**
50-year interval														
1693–1750	−0.49	−1.27	0.25	0.91	−0.21	−0.62	−1.12	−3.57***	0.34	0.90	0.35	1.02	0.21	1.21
1751–1800	−0.94	−3.81***	0.76	2.60***	0.33	1.02	−0.65	−1.86*	0.12	0.51	0.63	2.04**	0.27	1.70*
1801–50	−1.04	−1.91*	−1.21	−3.59***	−1.26	−2.82***	0.05	0.11	0.79	1.69*	1.19	2.44**	0.92	3.95***
1851–1900	−0.12	−0.50	−0.24	−1.02	0.42	2.19**	−0.02	−0.13	0.10	0.48	0.19	0.93	0.09	0.75
1901–50	−0.60	−2.27**	0.16	0.55	0.31	0.73	0.16	0.59	0.12	0.42	−0.23	−0.57	0.11	0.61
1951–2009	−0.01	−0.02	1.18	2.46**	−1.38	−2.79***	−0.34	−0.58	0.03	0.05	1.53	1.94*	1.05	3.09***

Table VI. Calendar month effects and the Halloween effect: Robust regressions

Table VI presents the coefficients estimates (percentage) and the Chi-square of the calendar month effect and the Halloween effect from the robust regressions of the equation: $r_t = \alpha + \beta_m D_{mt} + \varepsilon_t$, where r_t is the continuously compounded monthly returns, D_{mt} is the dummy variable of the calendar month m (or the Halloween dummy that equals 1, if the month falls on the period November through April and 0 otherwise for the Halloween effect regression). The robust regressions are based on M-estimation introduced by Huber (1973). The sample is subdivided into three subperiods of approximately 100-year intervals and six subperiods of 50-year intervals. ***denotes significance at the 1% level; **denotes significance at 5% level; *denotes significance at 10% level.

	January		February		March		April		May		June	
Sample period	β	X^2	β	X^2	β	X^2	β	X^2	β	X^2	β	X^2
1693–2009	0.26	3.34*	−0.15	1.11	−0.18	1.62	0.25	3.18	0.16	1.40	−0.22	2.48
100-year interval												
1693–1800	−0.56	7.52***	0.01	0.00	0.33	2.56	0.32	2.44	0.55	7.19***	−0.04	0.04
1801–1900	0.66	12.40***	0.13	0.47	−0.36	3.77*	−0.20	1.18	0.17	0.78	0.11	0.36
1901–2009	0.63	3.37*	−0.75	4.74**	−0.39	1.28	0.99	8.33***	−0.43	1.57	−1.10	10.33***
50-year interval												
1693–1750	−0.14	0.17	−0.18	0.31	−0.40	1.48	0.73	4.83**	0.93	8.00***	0.04	0.02
1751–1800	−0.92	13.20***	0.24	0.90	0.96	14.50***	−0.02	0.01	0.23	0.80	−0.13	0.26
1801–50	0.48	1.90	0.27	0.60	−0.37	1.17	−0.10	0.08	0.52	2.30	0.34	0.96
1851–1900	0.80	15.45***	−0.12	0.37	−0.41	4.06**	−0.31	2.22	−0.13	0.40	0.06	0.08
1901–50	0.75	5.72**	−0.50	2.48	−0.94	9.06***	0.33	1.08	0.16	0.27	−0.60	3.64*
1951–2009	0.46	0.56	−0.73	1.46	0.22	0.13	2.02	11.23***	−1.22	4.02**	−1.58	6.82***

	July		August		September		October		November		December		Halloween	
Sample period	β	X^2	β	X^2	β	X^2	β	X^2	β	X^2	β	X^2	β	X^2
1693–2009	−0.53	14.32***	0.26	3.48*	−0.08	0.36	−0.46	11.01***	0.25	3.19*	0.41	8.76***	0.26	11.16***
100-year interval														
1693–1800	−0.57	7.71***	0.56	7.57***	−0.21	1.04	−1.00	24.64***	0.12	0.33	0.48	5.41**	0.22	3.59*
1801–1900	−0.51	7.55***	−0.25	1.73	−0.15	0.62	−0.18	0.91	0.15	0.67	0.50	7.18***	0.25	5.90**
1901–2009	−0.26	0.56	0.68	3.97**	0.17	0.25	−0.09	0.07	0.44	1.63	0.18	0.27	0.32	2.86*
50-year interval														
1693–1750	−0.48	2.15	0.54	2.74*	−1.02	9.41***	−1.16	12.65***	0.41	1.54	0.49	2.26	0.29	2.46
1751–1800	−0.68	7.09***	0.58	5.18**	0.38	2.17	−0.91	12.96***	−0.21	0.66	0.51	3.93**	0.18	1.58
1801–50	−0.95	7.73***	−0.78	5.09**	−0.67	3.82*	−0.26	0.57	0.50	2.14	1.15	11.27***	0.58	9.45***
1851–1900	−0.17	0.71	0.07	0.10	0.26	1.62	−0.14	0.45	−0.07	0.10	0.20	0.99	0.01	0.01
1901–50	0.05	0.03	0.44	1.97	0.60	3.64*	−0.20	0.38	0.24	0.56	−0.47	2.26	−0.16	0.83
1951–2009	−0.65	1.13	0.90	2.20	−0.66	1.16	0.03	0.00	0.13	0.05	1.02	2.84*	0.98	8.46***

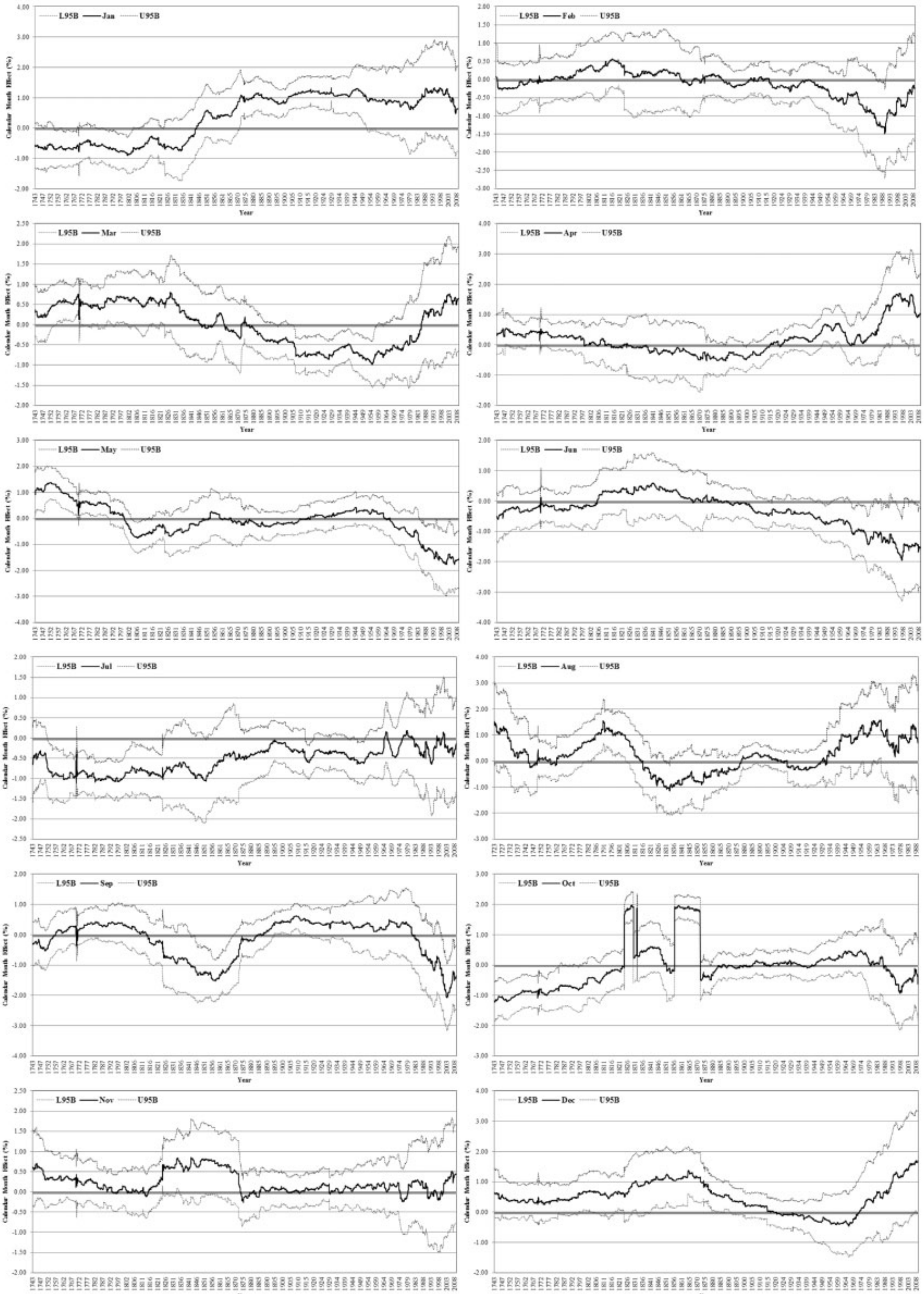

Figure 6. Estimates of 50-year rolling window regressions for the 12 calendar month effects and the Halloween effect estimated from time-varying volatility GARCH (1,1) models. The dark solid line indicates the coefficient estimates of the effect, the light dotted lines indicate the upper and the lower 95% bounds based on the GFD index.

Figure 6. Continued.

For the robust regression, we use the M-estimation introduced by Huber (1973), which is considered appropriate when the dependent variable may contain outliers.

Tables V and VI contain our estimation results for these regressions and Figures 6 and 7 contain the plots for the rolling window regressions.[20]

First, it may be good to note from the GARCH rolling windows and the OLS robust regressions that the widening of the confidence bounds seem to have disappeared and these tend to be the same size over time, suggesting that loss of power due to time-varying volatility is no longer an issue.

How far does this affect our results? If we use the same criteria, an overall significant effect and coefficients that must be of the correct sign for almost our full sample, when we use rolling windows of 50 years, we find based on the GARCH models that October drops out and positive November and December effects may resurface. Using OLS Robust regressions, we would probably reverse that conclusion and include October while again dropping November and December. The robustness tests seem to increase the strength for a July effect and a Halloween effect. The July effect is now significantly present in two out of 100-year subperiods. The Halloween effect is significant in all three 100-year subsamples. While both effects seem to be a bit stronger after we control for outliers and GARCH effects, there are still many 50-year periods when the effects are not significant.

5.2 VALUE WEIGHTED AND EQUALLY WEIGHTED INDICES

As Table I shows some indices are value weighted and others equally weighted. To determine whether this might affect our results, we try to

[20] The sudden shifts in the October GARCH plot seem to be caused by three subsequent months with high returns (November 1824–January 1825, with returns of 26.2%, 24.18%, and 53.53%, respectively). Once we remove these three observations, the shifts disappear. The GARCH model with *t*-distributed errors does not show the shifts.

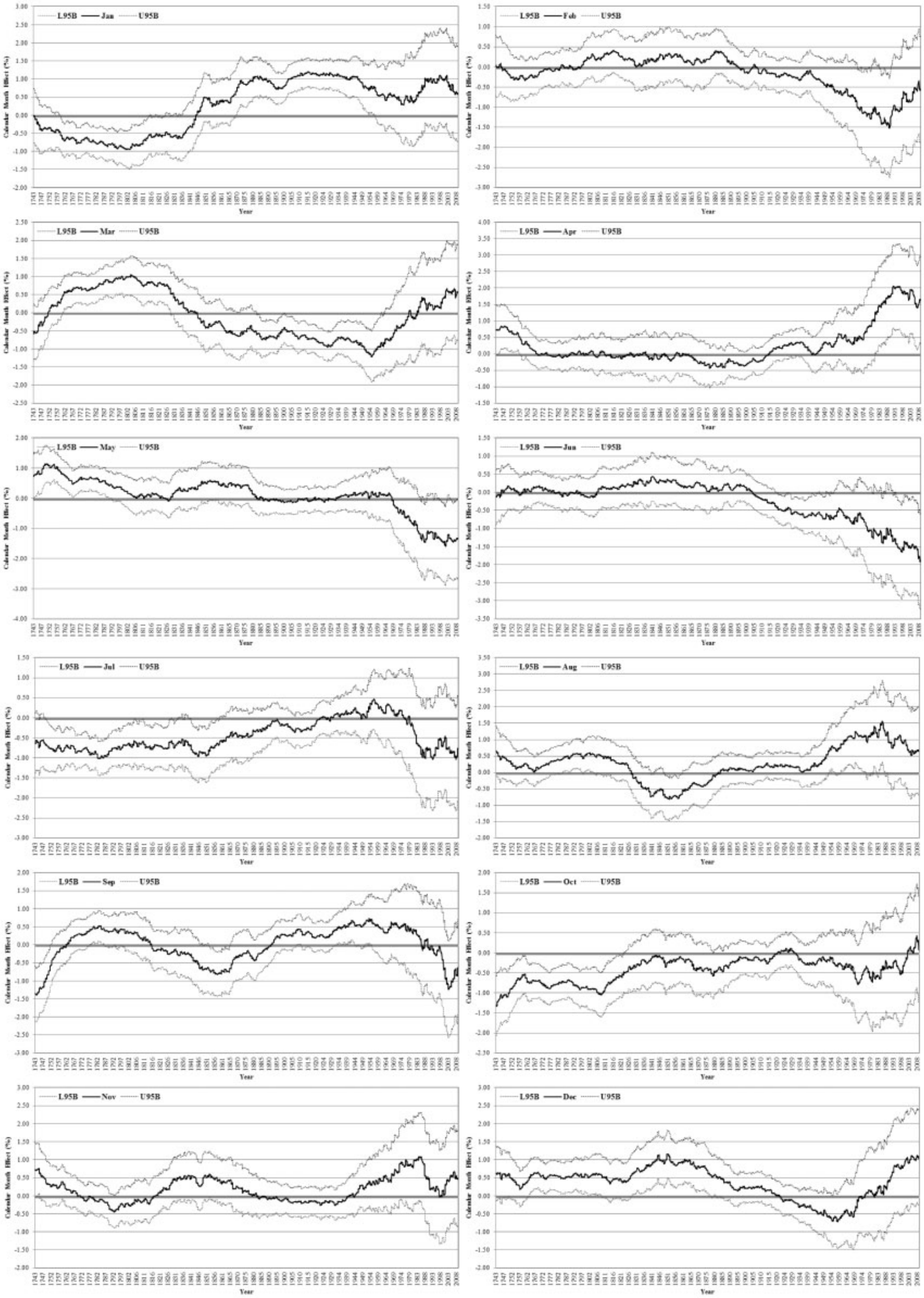

Figure 7. Estimates of the 50-year rolling window regressions of the 12 calendar month effects and the Halloween effect estimated from robust regressions based on M-estimation introduced in Huber (1973), the dark solid line indicates the coefficient estimates of the effect, the light dotted lines indicate the upper and the lower 95% bounds. Results are for the Gobal Financial Market index.

Figure 7. Continued.

construct a value-weighted index throughout. First, we construct a market value-weighted index for the three companies based on the individual price series for these three shares and calculate a value-weighted index assuming that there are no changes in the number of shares outstanding (*de facto* a price index). While there is some evidence (see, for instance, Shea, 2000) that these companies have issued shares and repurchased shares, these actions are infrequent and we have no exact details. So this is the closest we can get to a value-weighted index for this period. The other time period that uses an equally weighted index is from 1851 to 1906. Here, we were able to extend the Banker's magazine index backward to August 1887. For the period 1851–70, we use the value-weighted index constructed by Acheson *et al.* (2009). This leaves only a period of 16 years (1871–87) equally weighted.

Figure 8 shows if we replace the equally weighted parts with the value-weighted parts (apart from 1851 to 1887) and reestimate our results, these are hardly affected.

This is not surprising as the three stocks value-weighted index give almost similar results to the GFD index, because the market shares of these stocks were, with the exception of the South Sea Bubble in 1720, relatively stable over time.

5.3 DIVIDENDS

Dividends may influence these results, if large dividend payments cluster in specific months. In that case by using a price index, we could overestimate the significance of a negative effect, or underestimate a positive effect, in those months.

It is hard to conclusively determine whether there might be an effect over our full sample, but for two subsamples, we find little evidence that dividend clustering can explain our result. Thanks to a very thorough study of UK market returns and dividend payments by Dimson and Marsh (2001), we can conclude that in more recent periods the impact seems marginal and dominated by other differences in index construction. Dimson and Marsh

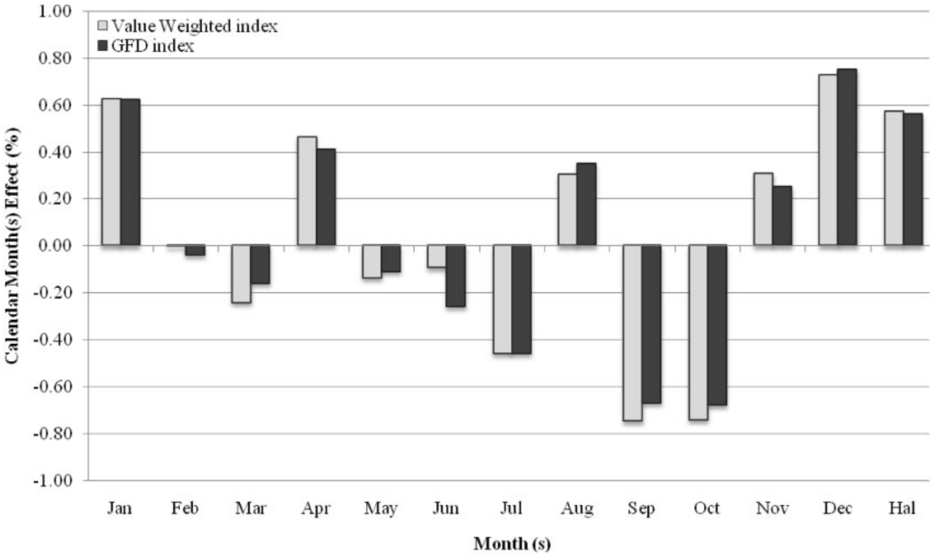

Figure 8. Calendar month(s) effect in GFD index and the constructed value-weighted index.

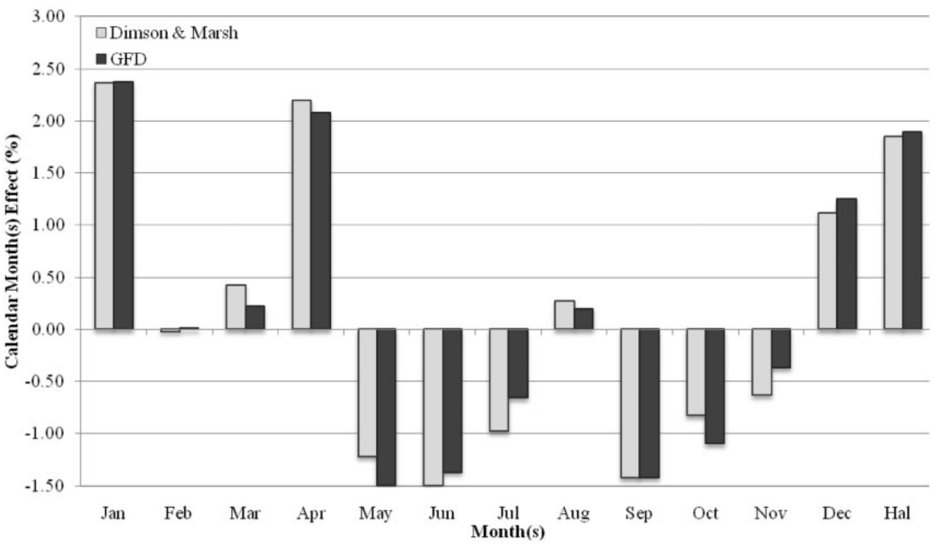

Figure 9. Calendar month(s) effect in the GFD index and the Dimson and Marsh (2001) index (1955–99).

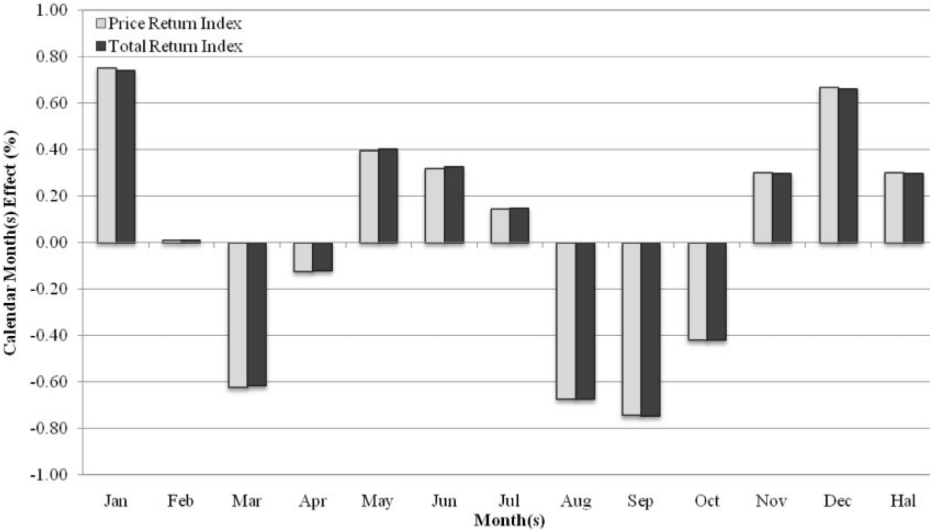

Figure 10. Calendar month(s) effect in UK price index and total return index data of Acheson *et al*. (2009) (1825–70).

(2001) construct an index, including dividends and report monthly UK equity premia from 1955 to 1999. In Figure 9, we compare their monthly excess returns (in deviation from the average of the other 11 months) with our data. To be consistent with Dimson and Marsh (2001), we also subtract 3-month UK Treasury bill yields from the returns of the GFD index.

These results are similar and do not seem to change our main results. This may not come as a surprise, because Dimson and Marsh (2001) also document that the largest difference between high and low dividend months is, at most, 9% of the total dividends.

Thanks to the extensive work of Acheson *et al*. (2009), we can make a more direct comparison over the 1825–70 period. They report monthly dividend yields. In Figure 10, we compare monthly seasonals based on both their value-weighted price index and their total return index, which includes dividends.

As the figure shows, annual dividends of 4.5% are almost equally distributed over the different months. A formal test also reveals no significant seasonalities in these dividend payments.

Unfortunately, for the other periods, exact evidence on the timing and the size of dividend payments is not available. Shea (2000) reports annual returns including dividends for the big three companies. If we combine these with the price information for these individual shares, we can extract the annual dividend yields. These are, on average, 6% up to 1,719, and then ranging between 4% and 5%. This suggests annual dividends up to 1,834 of ~5%. In our case, the distribution of these dividend payments over the months is more important. Neal (1987) documents that, just as in more recent history, UK dividends were paid semi-annually and different stocks would go ex-dividend in different months. For instance, The South Sea Company paid dividends in May and November, whereas the Bank of England paid dividends in March and September, as did the East Indies Company.[21]

Semi-annual dividend payments, at least, will not have an impact on the Halloween effect and, while the evidence suggests that the influence of dividends in the past should not be large, we cannot completely rule out that it may have an impact on the estimates over time. As, however, for most of our sample July and October have not been the dominant dividend months, this suggests that other months may have done relatively better than we document and, thus, that we underestimate the negative effects.

5.4 INTERACTION BETWEEN SEASONALS

With both July and October as consistently negative months, another question that might be raised is—that if one is willing to accept a Halloween effect exists—as to whether the negative returns in these 2 months may be the cause of the Halloween effect. Both months with negative returns on average fall in the summer period that is the poor performance period in the Halloween effect. To verify this, we reestimate the Halloween indicator regression controlling for both October and July. This reduces the Halloween effect marginally. Average monthly winter returns are 0.56% higher than the summer months without the two control variables (t-value of 4.04). If we include the dummies, then monthly winter returns are 0.44% higher (t-value of 3.16). The same conclusion holds if we include the overall significantly positive months (January, April, and December) jointly with a Halloween dummy. Halloween returns remain significant at 0.30% per month higher (t-value 2.05). The monthly July and October anomalies also remain significant, if we control for the overall significantly positive months. If we include January, April, and December dummies, the July and October effect still remains significantly negative, with −0.34% and −0.54% lower average returns on average (t-values −1.69 and −2.15, respectively).

[21] Private correspondence with Bryan Taylor of Global Financial Data.

6. Conclusion

We find that what should be a relatively simple question: whether or not there are seasonal monthly anomalies, strongly depends on the sample period considered. We show that many calendar months significantly outperform, or underperform, the market in our sample, but that few have done so persistently over the 300 years. This result confirms the potential problems caused by data snooping, noise and selection bias, and highlights the importance of studying long time series and suggests that many if not all calendar month anomalies may be spurious. Based on 50-year samples, it is hard to detect any persistent statistically significant anomalies. Our rolling window regressions show significant results fluctuate over time. For almost every month, we can find 50 years of fame. Conclusions vary strongly based on the selected sample size even over 100-year intervals. For example, the January effect switches from significantly negative to significantly positive based on the 100-year samples. If we only consider the full sample, we find 4 monthly anomalies robust across different estimation methods (significantly positive returns for January and December and significantly below average returns for July and October) and also a positive Halloween or Sell in May effect. However, in that case, one should be aware that in extremely long subperiods, the effect may be reversed and significantly so. Again the January effect is a clear example: significantly positive over the full sample but significantly negative in the first 100 years. Therefore, whether or not, and which of these monthly anomalies exist, seems to depend strongly on sample periods and criteria we apply. Or in other words, these monthly anomalies may be in the eye of the beholder.

References

Acheson, G. G., Hickson, C. R., Turner, J. D., and Ye, Q. (2009) Rule Britannia! British stock market return, 1825-1870, *Journal of Economic History* **69**, 1107–1137.

Anderson, L. R., Gerlach, J. R., and DiTraglia, F. J. (2007) Yes, Wall Street, There is a January Effect! Evidence from laboratory auctions, *Journal of Behavioral Finance* **8**, 1–8.

Andrade, S. C., Chhaochharia, V., and Fuerst, M. E. (2012) "Sell in May and Go Away"Just Won't Go Away. Unpublished working paper, University of Miami.

Ariel, R. A. (1987) A monthly effect in stock returns, *Journal of Financial Economics* **18**, 161–174.

Barry, C. B. and Brown, S. J. (1984) Differential information and the small firm effect, *Journal of Financial Economics* **13**, 283–294.

Berges, A., McConnell, J. J., and Schlarbaum, G. G. (1984) The turn-of-the-year in Canada, *Journal of Finance* **39**, 185–192.

Bouman, S. and Jacobsen, B. (2002) The Halloween indicator, "Sell in May and go away": another puzzle, *American Economic Review* **92**, 1618–1635.

Brown, R. L. and Easton, S. A. (1989) Weak-form efficiency in the nineteenth century: a study of daily prices in London market for 3 per cent Consols, 1821-1860, *Economica* **56**, 61–70.

Brown, W. O. Jr, Burdekin, R. C. K., and Weidenmier, M. D. (2006) Volatility in an era of reduced uncertainty: lessons from Pax Britannica, *Journal of Financial Economics* **79**, 693–707.

Brown, P., Keim, D., Kleidon, A., and Marsh, T. A. (1982) Stock return seasonalities and the tax-loss selling hypothesis: analysis of the arguments and Australian evidence, *Journal of Financial Economics* **12**, 105–127.

Cao, M. and Wei, J. (2005) Stock market returns: a note on temperature anomaly, *Journal of Banking and Finance* **29**, 1559–1573.

Choudhry, T. (2001) Month of the year effect and January effect in pre-WW1 stock returns: evidence from a non-linear Garch model, *International Journal of Finance and Economics* **6**, 1–11.

Ciccone, S. J. (2011) Investor optimism, false hopes and the January effect, *Journal of Behavioral Finance* **12**, 158–168.

Clare, A. D., Psaradakis, Z., and Thomas, S. H. (1995) An analysis of seasonality in the UK equity market, *Economic Journal* **105**, 398–409.

Cooper, M. J., McConnell, J. J., and Ovtchinnikov, A. V. (2006) The other January effect, *Journal of Financial Economics* **82**, 315–341.

Corhay, A., Hawawini, G., and Michel, P. (1987) Seasonality in the risk-return relationship: some international evidence, *Journal of Finance* **42**, 49–68.

Darrat, A. F., Li, B., Liu, B., and Su, J. J. (2011) A fresh look at seasonal anomalies: an international perspective, *International Journal of Business and Economics* **10**, 93–116.

Dimson, E. and Marsh, P. (2001) U.K. Financial Market Returns, 1955-2000, *Journal of Business* **74**, 1–31.

Fountas, S. and Segredakis, K. N. (2002) Emerging stock markets return seasonalities: the January effect and the tax-loss selling hypothesis, *Applied Financial Economics* **12**, 291–299.

French, K. R. (1980) Stock returns and the weekend effect, *Journal of Financial Economics* **8**, 55–69.

Gayer, A. D., Rostow, W. W., and Schwartz, A. J. (1975) *The Growth and Fluctuation of the British Economy 1790-1850: An Historical, Statistical, and Theoretical Study of Britain's Economic Development*, The Harvester Press Limited, Hassocks, England.

Gerlach, J. R. (2007) Macroeconomic news and stock market calendar and wealther anomalies, *Journal of Financial Research* 283–300.

Gibbons, M. R. and Hess, P. (1981) Day of the week effects and asset returns, *Journal of Business* **54**, 579–596.

Glasner, D. (1997) *Business Cycles and Depressions – An Encyclopedia*, Garland Publishing Inc, New York & London.

Goetzmann, W. N. (1993a) Accounting for taste: art and the financial market over three centuries, *American Economic Review* **83**, 1370–1376.

Goetzmann, W. N. (1993b) Patterns in three centuries of stock market prices, *Journal of Business* **66**, 249–270.

Goetzmann, W. N. and Ibbotson, R. G. (2005) History and the equity risk premium, in: R. Mehra (ed.), *Handbook of the Equity Risk Premium*, Elsevier, Amsterdam, pp. 515–529.

Goetzmann, W. N., Ibbotson, R. G., and Peng, L. (2001) A new historical database for the NYSE 1815 to 1925: performance and predictability, *Journal of Financial Markets* **4**, 1–32.

Goetzmann, W. N. and Jorion, P. (1995) A longer look at dividend yields, *Journal of Business* **68**, 483–508.

Goetzmann, W. N., Li, L., and Rouwenhorst, K. G. (2005) Long-term global market correlations, *Journal of Business* **78**, 1–38.

Grossman, R. S. (2002) New indices of British equity prices, 1870-1913, *Journal of Economic History* **62**, 121–146.

Grossman, R. S. and Shore, S. H. (2006) The cross section of stock returns before World War I, *Journal of Financial and Quantitative Analysis* **41**, 271–294.

Gultekin, M. N. and Gultekin, N. B. (1983) Stock market seasonality international evidence, *Journal of Financial Economics* **12**, 469–481.

Harrison, P. (1998) Similarities in the Distribution of Stock Market Price Changes between the Eighteenth and Twentieth Centuries, *Journal of Business* **71**, 55–79.

Haugen, R. and Lakonishok, J. (1987) *The Incredible January Effect: The Stock Market's Unsolved Mystery*, Dow-Jones-Irwin, Homewood.

Heston, S. L. and Sadka, R. (2007) Seasonality in the cross-section of stock returns, *Journal of Financial Economics* **87**, 418–445.

Ho, Y.-K. (1990) Stock return seasonalities in Asia Pacific markets, *Journal of International Financial Management and Accounting* **2**, 48–77.

Hong, H. and Yu, J. (2009) Gone fishin': seasonality in trading activity and asset prices, *Journal of Financial Markets* **12**, 672–702.

Huber, P. J. (1973) Robust regression: asymptotics, Conjectures, and Monte Carlo, *Annals of Statistics* **1**, 799–821.

Jacobsen, B. and Dannenburg, D. (2003) Volatility clustering in monthly stock returns, *Journal of Empirical Finance* **10**, 479–503.

Jacobsen, B., Mamun, A., and Visaltanachoti, N. (2005) Seasonal, size and value anomalies. Unpublished working paper, Massey University, University of Saskatchewan.

Jacobsen, B. and Marquering, W. (2008) Is it the weather? *Journal of Banking and Finance* **32**, 526–540.

Jacobsen, B. and Visaltanachoti, N. (2009) The Halloween Effect in US Sectors, *Financial Review* **44**, 437–459.

Jones, S. L., Lee, W., and Apenbrink, R. (1991) New evidence on the January effect before personal income tax, *Journal of Finance* **46**, 1909–1924.

Jones, C. P., Pearce, D. K., and Wilson, J. W. (1987) Can tax-loss selling explain the January effect? A note, *Journal of Finance* **42**, 453–461.

Jorion, P. and Goetzmann, W. N. (1999) Global stock markets in the twentieth century, *Journal of Finance* **54**, 953–980.

Kamstra, M. J., Kramer, L. A., and Levi, M. D. (2003) Winter blues: a sad stock market cycle, *American Economic Review* **93**, 324–343.

Kato, K. and Schallheim, J. S. (1985) Seasonal and size anomalies in the Japanese stock market, *Journal of Financial and Quantitative Analysis* **20**, 243–260.

Keim, D. (1983) Size-related anomalies and stock return seasonality: further empirical evidence, *Journal of Financial Economics* **12**, 13–32.

Kelly, P.J. and Meschke, F. (2010) Sentiment and stock returns: the SAD anomaly revisited, *Journal of Banking and Finance* **34**, 1308–1326.

Kim, D. (2006) On the information uncertainty risk and the January effect, *Journal of Business* **79**, 2127–2161.

Lakonishok, J. and Levi, M. D. (1982) Weekend effects on stock returns, *Journal of Finance* **37**, 883–889.

Lakonishok, J. and Smidt, S. (1988) Are seasonal anomalies real? A ninety-year perspective, *Review of Financial Studies* **1**, 403–425.

Lakonishok, J., Shleifer, A., Thaler, R., and Vishny, R. (1991) Window dressing by pension fund managers, *American Economic Review* **81**, 227–231.

Lean, H. H. (2011) The Halloween puzzle in selected Asian stock markets, *International Journal of Economics and Management* **5**, 216–225.

Lundblad, C. (2007) The risk return tradeoff in the long run: 1836-2003, *Journal of Financial Economics* **85**, 125–150.

Merton, R. C. (1980) On estimating the expected return on the market, *Journal of Financial Economics* **8**, 323–361.

Mirowski, P. (1981) The rise (and retreat) of a market: English joint stock shares in the eighteenth century, *Journal of Economic History* **41**, 559–577.

Neal, L. (1987) The integration and efficiency of the London and Amsterdam stock markets in the eighteenth century, *Journal of Economic History* **47**, 97–115.

Newey, W. K. and West, K. D. (1987) A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* **55**, 7003–7708.

Ng, L. and Wang, Q. (2004) Institutional trading and the turn-of-the-year effect, *Journal of Financial Economics* **74**, 343–366.

Ogden, J. P. (1990) Turn-of-month evaluations of liquid profits and stock returns: a common explanation for the monthly and January effects, *Journal of Finance* **45**, 1259–1272.

Ogden, J. P. (2003) The calendar structure of risk and expected returns on stocks and bonds, *Journal of Financial Economics* **70**, 29–67.

Ogden, J. P. and Fitzpatrick, J. (2010) Do five asset pricing anomalies share a common mispricing factors? Multifaceted empirical analyses of failure risk proxies, external financing, and stock returns. Unpublished working paper, SUNY at Buffalo.

Poterba, J. M. and Weisbenner, S. J. (2001) Capital gains tax rules, tax-loss trading, and turn-of-the-year returns, *Journal of Finance* **56**, 353–368.

Reinganum, M. R. (1983) The anomalous stock market behavior of small firms in January - empirical tests for tax-loss selling effects, *Journal of Financial Economics* **12**, 89–104.

Reinganum, M. R. and Shapiro, A. C. (1987) Taxes and stock return seasonality: evidence from the London stock exchange, *Journal of Business* **60**, 281–295.

Roll, R. (1983) Vas ist das? The turn-of-the-year effect and the return premia of small firms, *Journal of Portfolio Management* **Winter**, 18–28.

Rozeff, M. S. and Kinney, W. R. (1976) Capital market seasonality: the case of stock returns, *Journal of Financial Economics* **3**, 379–402.

Schultz, P. (1985) Personal income taxes and the January effect: small firm stock returns before the War Revenue Act of 1917: a note, *Journal of Finance* **40**, 333–343.

Shea, S. G. (2000) The course of the exchange: measuring and interpreting returns process in 18th and early 19th century Britain. Working paper, Department of Economics University of St Andrews.

Shiller, R. J. (1989) Comovements in stock prices and comovements in dividends, *Journal of Finance* **44**, 719–729.

Starks, L. T., Yong, L., and Zheng, L. (2006) Tax-loss selling and the January effect: evidence from municipal bond closed-end funds, *Journal of Finance* **LXI**, 3049–3067.

Swinkels, L. A. P. and Van Vliet, P. (2010) An anatomy of calendar effects. Unpublished working paper, Robeco Investment Solutions & Research, Robeco Asset Management.

Tinic, S. M., Adeshi, B. G., and West, R. R. (1987) Seasonality in Canadian stock prices: a test of the “tax-loss-selling” hypothesis, *Journal of Financial and Quantitative Analysis* **22**, 51–63.

Wachtel, S. B. (1942) Certain observations on seasonal movements in stock prices, *Journal of Business* **15**, 184–193.

Wilson, J. W. and Jones, C. P. (2002) An anaysis of the S&P500 index and Cowles’ extensions: price indexes and stock returns, *Journal of Business* **75**, 505–533.

Van den Bergh, W. M. and Wessels, R. E. (1985) Stock market seasonality and taxes: an examination of the tax-loss selling hypothesis, *Journal of Business Finance & Accounting* **12**, 515–530.

Van der Gugten, T. (2010) Stock market calendar anomalies and macroeconomic news announcements. Unpublished working paper, Erasmus University Rotterdam.

Zarour, B. A. (2007) The Halloween effect anomaly: evidence from some Arab countries equity markets, *Studies in Business and Economics* **13**, 68–76.