

# Risk and ethical investment: Empirical evidence from Dow Jones Islamic indexes

Amélie Charles, Olivier Darné, Adrian Pop

► **To cite this version:**

Amélie Charles, Olivier Darné, Adrian Pop. Risk and ethical investment: Empirical evidence from Dow Jones Islamic indexes. *Research in International Business and Finance*, Elsevier, 2015, 35, pp.33-56. <10.1016/j.ribaf.2015.03.003>. <hal-01153899>

**HAL Id: hal-01153899**

**<http://hal-audencia.archives-ouvertes.fr/hal-01153899>**

Submitted on 20 May 2015

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



# **Risk and Ethical Investment: Empirical Evidence from Dow Jones Islamic Indexes**

**Amélie CHARLES**

Audencia Nantes, School of Management

**Olivier DARNÉ**

LEMNA, University of Nantes

**Adrian POP**

LEMNA, University of Nantes

## **Abstract**

This paper contributes to the literature on the impact of the *Shari'ah* filtering criteria on the risk of Dow Jones Islamic indexes relative to their conventional counterparts. We show that Islamic and conventional indexes are affected by the same extreme events which can bias the estimation of the risk, especially the period of the Global Financial Crisis of 2007-2008 and its aftermath which is characterized by a very high level of volatility. Then, we examine whether the Islamic indexes are more risky than the conventional indexes using different risk measures. We also analyzed the performance of both indexes from various risk-adjusted performance measures. Overall, the Islamic indexes seem to be more risky than their conventional counterparts as well as exhibit a higher performance on the full period (1996-2013). The sample period is further divided into low volatility period and high volatility period based on the detection of structural breaks in the volatility. The results also show that both indexes have been affected by variance changes. We show that most of the Islamic indexes have higher level of risk than the conventional indexes, whatever the sub-periods. Consequently, this finding means that the Islamic indexes are riskier than the non-Islamic indexes. We also find that in most cases the Islamic indexes either outperform the non-Islamic indexes or there is no significant difference in performance between both indexes. These findings can be explained as a consequence of less diversification in Islamic indexes, leading to higher concentration risk in some sectors, such as basic material, industrial and technology firms. Further, we also find

some differences of risk and performance between the jurisdictions.

*Keywords:* Risk, Performance, Islamic indexes, Conventional indexes.

*JEL Classification:* G01, G21, G28, C12, C14.

# 1 Introduction

There has been large-scale growth in Islamic finance and banking in Muslim countries and around the world during the last twenty years. This growth is influenced by factors including the introduction of broad macroeconomic and structural reforms in financial systems, the liberalization of capital movements, privatization, the global integration of financial markets, and the introduction of innovative and new Islamic products (Zaher and Hassan, 2001). Nevertheless, Islamic finance has entered a bright new stage of development, emerging after the global financial crisis as a more equitable and efficient alternative to the Western approach (Ahmed, 2010; Alasrag, 2010).

Islamic finance, as is well known, is based on the application of classical Islamic law in the management of money: this implies the prohibition of interest, of excessive risk, of gambling, the exclusion of investments in arms, alcohol, casinos, tobacco, pornography and pork, and a major attention on social welfare. Like any other modern avatars of ethical investments<sup>1</sup>, such as *green*, *faith* or *socially responsible* investments, the Islamic investing aims at generating low volatility returns and value enhancement opportunities by focusing on low-debt, non-financial, social-ethical investment vehicles (De Lorenzo, 2001).<sup>2</sup>

Most defenders of ethical and Islamic investing point out that the underlying screening process implies that the eligible selected firms have stronger and much stable financial positions and are more successful than the excluded firms. On the other hand, because of monitoring and screening costs or potentially reduced levels of diversification inherent to the screening process, unscreened firms may outperform ethical or Islamic investment vehicles (Temper, 1991; Sauer, 1997; Iqbal and Llewellyn, 2002; Hussein, 2004). Moreover, as noted by Hussein and Omran (2005), the screening process tends to systematically exclude large firms from the global universe of investable equities, which implies that the remaining eligible firms are smaller and exhibit more volatile returns.

Langbein and Posner (1980) are the first to argue that ethical investments may involve higher risk than their conventional counterparts. However, by comparing the returns of a well-established index of socially responsible firms (Domini Social Index, DSI) to other conventional market portfolios (such as

---

<sup>1</sup>Cowton (1994) defines ethical investment as the use of ethical and social criteria in the selection and management of investment portfolios, generally consisting of company shares. Ethical investors are not only concerned about the financial returns on their portfolios and the risks involved but also with the characteristics of the companies in which the funds are placed.

<sup>2</sup>The most important difference between Islamic and other ethical funds is that in addition to the exclusion of particular sectors, Islamic funds do not deal in fixed income market and the receipt and payment of interest is not permitted (Hussein, 2004).

S&P 500 or CRSP value-weighted market index), Sauer (1997) fails to find that ethical screening implies higher volatility of returns and reduced financial performance. In the same vein, Statman (2000) finds that the risk-adjusted returns of S&P 500 over the 1990–1998 period do not differ significantly from those of the DSI, although the ethical index exhibits a slightly higher volatility than its conventional counterpart. More recently, Al-Zoubi and Maghyereh (2007) focus on the relative risk performance of Islamic finance while comparing the risk performance of the DJ Islamic index with the DJ World index over the period 1996-2005. They find that the risk level of Islamic index is significantly lower than the market basket of stocks.

The present paper contributes to the literature on the impact of the *Shari'ah* filtering criteria on the risk of Dow Jones Islamic indexes relative to their conventional counterparts. Understanding the behavior of risk in Islamic and conventional indexes is important to risk management, derivative pricing and hedging, market making, market timing, portfolio selection, and many other financial activities. Contrary to the previous studies that only focused on performance measures, in this study we examine whether Islamic indexes are more risky than non-Islamic indexes. The studies on risk of ethical indexes are very sparse. However, this is beneficial to investors since risk is one of the main characteristic to formulate a good investment portfolio. Thus, in this paper, we will try to response to the following question: Are the Islamic indexes more risky than the conventional indexes?

Large daily swings that are neither unique, unusual nor as dramatic have a substantial impact on stock market returns (see, e.g., Wang et al., 2009). Events such as wars, terrorism, and bankruptcy are known as “black swans”.<sup>3</sup> Failing to take explicit account of the fact that such extraordinary movements have occurred in the past – and will occur in the future – is therefore a serious omission (Friedman and Laibson, 1989). For example, Bloom (2009) shows that uncertainty dramatically increases after major economic and political shocks. One way to identify black swans or (infrequent) large shocks is intervention analysis, introduced by Box and Tiao (1975) to attempt to statistically appraise these types of shocks (or outliers). Intervention analysis is used to assess the impact of a known or unknown event on the time series. The main focus is to estimate the effect of such events on the series. Intervention analysis forms the basis for many outlier modelling procedures. Outliers are usually present in financial

---

<sup>3</sup>According to Taleb (2007), a black swan is an event with three attributes: (i) It is an outlier, lying outside the realm of regular expectations because nothing in the past can convincingly point to its occurrence; (ii) it carries an extreme impact; and (iii) despite being an outlier, plausible explanations for its occurrence can be found after the fact, thus implying that it is explainable and predictable.

time series (see, e.g., Charles and Darné, 2005; Bali and Guirguis, 2007). Further, it is well known that the outliers bias the estimation of the variance, and thus the estimation of the risk (see, e.g., Charles, 2008; Carnero et al., 2007; Charles and Darné; 2014).

Firstly, in this paper, we employ the semi-parametric procedure to detect additive outliers proposed by Laurent, Lecourt and Palm (LLP) (2014) based on the GJR model of Glosten, Jagannathan, and Runkle (1993) that accounts for the so-called leverage effect<sup>4</sup> on several major (conventional and Islamic) daily stock market indexes spanning January 1<sup>st</sup>, 1996 to March 15<sup>th</sup>, 2013. We also examine whether these outliers tend to be social, political or economic, and whether the Islamic indexes are more affected by these extreme events than the conventional indexes. We show that Islamic and conventional indexes are affected by the same events in most cases, especially the period of the Global Financial Crisis (GFC) of 2007-2008 and its aftermath as well as events associated with the 1997 Asian financial crisis, the Russian crisis of summer 1998, the terrorist attacks in the US on September 11th, 2001, and other macroeconomic news. Then, we apply traditional (standard deviation and semivariance) and standard (Value-at-Risk and expected shortfall) risk measures on the original and adjusted-outlier returns of both indexes. The results show that the risk measures are higher for the Islamic indexes than for their conventional counterparts, suggesting that the Islamic indexes are more risky. Further, the adjusted-outlier data display a lower level of risk than the original data. We also analyzed the performance of both indexes from various risk-adjusted performance measures, namely Sharpe ratio, Treynor ratio, Jensen's alpha, Black-Treynor ratio, Sortino ratio and Reward-to-VaR ratio. The results show that in most cases the Islamic indexes seem to exhibit higher risk-adjusted performance than their conventional counterpart.

Secondly, we use an appropriate methodology to identify breakpoints and sudden shifts in volatility to define low volatility periods and high volatility periods. A relatively recent approach to test for volatility shifts is the iterative cumulative sums of squares (ICSS) algorithm (Inclán and Tiao, 1994; Sansó et al., 2004). This algorithm allows for detecting multiple breakpoints in variance and has been extensively used for identifying changes in the volatility of financial time series (Hammoumdeh and Li, 2008; Kasman, 2009; and Wang and Moore, 2009, among others). Nevertheless, Rodrigues and Rubia (2011) show that the asymptotic distribution of the ICSS test statistics varies under additive outliers.<sup>5</sup> Therefore, we then apply the ICSS algorithm to detect sudden changes in volatility of both indexes using

---

<sup>4</sup>Stock returns exhibit some degree of asymmetry in their conditional variances, i.e. that market participants overreact to bad news as compared to good news (Black, 1976; French, Schwert and Stambaugh, 1987; Bollerslev, Chow, and Kroner, 1992). Albaity and Ahmad (2011a, 2011b) found evidence of leverage effect in the Dow Jones Islamic Market Index.

<sup>5</sup>An additive outlier causes an immediate and one-shot effect on the observed series.

the outlier-corrected return series. The results show that both Islamic and conventional indexes have been affected by variance changes, especially the period following the 2008 financial crisis (2008-2010) which is associated with very high levels of volatility. We then apply the measures of risk between both indexes from the different sub-periods. We show that most of the Islamic indexes have higher level of risk than the conventional indexes, whatever the sub-periods. Consequently, this finding means that the Islamic indexes are riskier than the non-Islamic indexes. We also find that in most cases the Islamic indexes either outperform the non-Islamic indexes or there is no significant difference in performance between both indexes. These findings can be explained as a consequence of less diversification in Islamic indexes, leading to higher concentration risk in some sectors, such as basic material, industrial and technology firms. Further, we also find some differences of risk and performance between the jurisdictions.

This article is organized as follows. Section 2 presents a brief survey of past empirical studies on performance of Islamic investing. Section 3 describes the sequential procedure for detecting outliers in stock market indexes returns, and the modified ICSS algorithm used to identify sudden variance breaks in Islamic and conventional indexes. The empirical results are discussed in Section 4. Finally, Section 5 concludes.

## **2 Review of literature**

Despite the increasing importance and popularity of performance of ethical investments over the past several years, the existing literature on Islamic investing contains only a few empirical studies (see Tables 1 and 2).<sup>6</sup>

Some studies investigate the impact of ethical screening on the performance of Islamic indexes relative to their conventional counterparts. Ahmad and Ibrahim (2002) and Albaity and Ahmad (2008) study Kuala Lumpur Shariah index (KLSI) with Kuala Lumpur Composite index (KLCI) in Malaysia during the period April 1999 to January 2002 and April 1999 to December 2005, respectively. Using various measures of risk-adjusted performance, they find that there are no significant difference in the risk-adjusted performance of both indices. Similar results are obtained by Hussein (2004) for the FTSE Global Islamic index (FTSEGII) during the period July 1996 to August 2003, Hussein (2007) for the FTSEGII and the DJ Islamic Market index (DJIM) over the period December 1996 to December 2004, Girard and Hassan (2008) for five FTSE Islamic indexes (FTSEGII and 4 regional indexes) from

---

<sup>6</sup>See Hassan and Girard (2011) for a review of empirical literature on faith-based Islamic investing.

December 1998 to December 2006, Hassan and Girard (2011) for seven indexes chosen from the Dow Jones Islamic Market (DJIM) Index over the period period January 1996 to November 2005, Dharani and Natarajan (2011a) for the Nifty Shariah index in India during the period January 2007 to December 2010, and Abbes (2012) on 35 MSCI Islamic stock market indices over the period of June 2002 to April 2012. All these studies indicate that there is no significant difference in performance of both indexes over the entire period, suggesting that the Islamic and conventional indexes are of similar performance. Consequently, it means that whether an investor invests in screened or non-screened index, the returns will not differ. Nevertheless, Hussein (2004, 2007), Hassan and Girard (2011) and Abbes (2012) find that the Islamic indices yield statistically significant positive abnormal returns in the bull market period, although it underperforms the counterpart indices in the bear market period.<sup>7</sup> This finding can guide investors in their investment decision by providing information on the risk and return relation during bull and bear periods. Further, Hussein (2007) also analyzes the long run performance with buy-and-hold returns and cumulative returns and shows that the Islamic index has a superior performance compared with its counterpart in the entire and first bull market (1996-2000) periods. On the other hand, the Islamic index fails to sustain its better performance over the bear and second bull market (2002-2004) periods. Hussein and Omran (2005) obtained the same results on the long run performance by comparing DJIM and its 13 subindexes with their DJ conventional indexes for the time period December 1995 through June 2003.

However, from a recent period, Ashraf and Mohammad (2014) show that Islamic indices, in general, perform better than conventional indices, by comparing the performance of 12 global and regional Islamic indices with conventional indices from June 2002 to May 2012. This result can be explained by the crisis periods. Ho et al. (2014) find that Islamic indices outperformed their conventional counterparts during crisis periods, Dotcom crisis (2000-2002) and Global Financial Crisis (2007-2008) periods, but results are inconclusive for the non-crisis periods, by comparing risk-adjusted performance of 12 global conventional and Islamic indices during the period 2000 to 2011. Jawadi et al. (2014) also show that Islamic indices appear to outperform conventional indices, particularly during the Global Financial Crisis,<sup>8</sup> while non-Islamic indices seem to be preferred in calm periods, by examining three regions (the World, the USA and Europe) from January 2000 through June 2011. Nevertheless, Abbes (2012) shows

---

<sup>7</sup>Note that Hussein (2004, 2007) and Hassan and Girard (2011) define the same period 1996-2000 as a bullish period, but there is some differences for the bearish period: 2000-2003 for Hussein (2004), 2000-2002 for Hussein (2007), and 2000-2005 for Hassan and Girard (2011).

<sup>8</sup>Note that the Global Financial Crisis is defined from 2007 to 2011 for Jawadi et al. (2014) and from 2007 to 2008 for Ho et al. (2014).



that in the crisis period (2008-2009) there is no difference for 35 MSCI Islamic stock market indices over the period of June 2002 to April 2012.

Moreover, Albaity and Ahmad (2011) compare the performance of two Islamic indexes, FTSE Global Islamic index and DJIM index, with two socially responsible indexes, FTSE For Good index and DJ Socially responsible world index, and find no significant difference between 1999 and 2007. Further, Al-Khazali et al. (2014) compare the performance of 9 DJ Islamic indexes to their DJ conventional counterparts using stochastic dominance analysis from January 1996 to December 2012. They find that all conventional indexes stochastically dominate Islamic indexes at second and third orders in all markets except the European market. Recently, Abdelsalam et al. (2014) analyze the performance persistence of Islamic and Socially Responsible Investment mutual funds from 2001 to 2011, using partial frontiers' approaches. Their results indicate that, for both types of funds, performance persistence actually exists, but only for the worst and, most notably, best funds.

Hakim and Rashidan (2002) find that the DJ Islamic indexes from 1999 to 2002 are influenced by other factors than the broad market movements in stock prices and interest rates from a cointegration analysis.<sup>9</sup> More importantly, their findings suggest that the Islamic index exhibits unique risk-returns characteristics and the screening process applied to exclude *Shari'ah*-non-compliant firms does not adversely affect the performance of the Islamic index. According to Sadegi (2008) the introduction of Bursa Malaysia Islamic index has positive and strong impact on the financial performance of the Shariah compliant stock from event study methodology.<sup>10</sup>

Finally, Al-Zoubi and Maghyreh (2007) focus on the relative risk performance of Islamic finance while comparing the risk performance of the DJ Islamic index with the DJ World index over the period 1996-2005. They find that the risk level of Islamic index is significantly lower than the market basket of stocks, using Value-at-Risk methodology.

---

<sup>9</sup>Yusof and Majid (2007) examine the extent to which the conditional volatilities of both Islamic and conventional indexes in Malaysia, KLCI and Rashid Hussain Berhad Islamic index, respectively, are affected by monetary policy variables. The most important finding is that the interest rate volatility affects the conventional, but not the Islamic, stock market volatility, which is quite intuitive. Recently, Shamsuddin (2014) finds that the DJIM index is immune to interest rate risk over the period January 1996 to April 2011. However, at the sectoral level some Islamic equity portfolios demonstrate exposure to interest rate risk. Overall, evidence of interest rate risk exposure is less pronounced among Islamic sector portfolios than that of their mainstream counterparts - the DJ World sector indices.

<sup>10</sup>Some studies study the causality and cointegration relationship between Islamic and non-Islamic indexes (e.g., Albaity and Ahmad, 2008; Biek and Wardhana, 2009; Ajmi et al., 2014).

Table 1: Selected studies on the performance of Islamic indexes.

Studies	Data	Sample	Methodologies
Ahmad and Ibrahim (2002)	KLSI vs KLCI	1999-2002 (M)	Sharpe ratio, Jensen's alpha, Treynor ratio
Hussein (2004)	FTSEGII vs FTSEWI	1996-2003 (M)	Sharpe ratio, Jensen's alpha, Treynor ratio
Hussein and Omran (2005)	13 DJ Islamic indexes vs their conventional indexes	1995-2003 (M)	Sharpe ratio, Jensen's alpha, Treynor ratio
Hussein (2007)	FTSEGII and DJIM vs FTSEWI and DJWI	1996-2004 (M)	Jensen's alpha
Al-Zoubi and Maghyereh (2007)	DJIM vs DJWI	1996-2005 (M)	Value-at-Risk methodology
Girard and Hassan (2008)	5 FTSE Islamic indexes vs their conventional indexes	1998-2006 (M)	Sharpe ratio, Jensen's alpha, Treynor ratio, 4-factors CAPM
Albaity and Ahmad (2008)	KLSI vs KLCI	1999-2005 (M)	Sharpe ratio, Jensen's alpha, Treynor ratio, cointegration and causality
Sadegi (2008)	KLSI		event study methodology
Dharani and Natarajan (2011)	Nifty Shariah index vs Nifty index	2007-2010 (M)	Sharpe ratio, Jensen's alpha, Treynor ratio
Hassan and Girard (2011)	7 DJ Islamic indexes vs their conventional indexes	1996-2005 (M)	Sharpe ratio, Jensen's alpha, Treynor ratio, 4-factors CAPM
Albaity and Ahmad (2011a)	2 Islamic indexes (FTSEGII and DJIM) 2 socially responsible indexes (FTSE4G and DJSRWI)	1999-2007 (M)	mean and GARCH models

Notes: D: daily; M: monthly. KLSI: Kuala Lumpur Shariah index; KLCI: Kuala Lumpur Composite index; FTSEGII: FTSE Global Islamic index; FTSEWI: FTSE All-World index; DJIM: DJ Islamic Market index; DJWI: DJ World index; FTSE4G: FTSE For Good index; DJSRWI: DJ Socially Responsible World index.

### 3 Methodology

#### 3.1 Outlier detection in GARCH models

Several studies have showed that financial data may be affected by contaminated observations (Balke and Fomby, 1994; Charles and Darné, 2005). This type of observations, called outliers, reflects extraordinary, infrequently occurring events or shocks that have important effects on macroeconomic and financial time series. There are several methods for detecting outliers in nonlinear setting, such as the method for additive jumps detection proposed by Franses and Ghijsels (1999) and Grané and Veiga

Table 2: Selected studies on the performance of Islamic indexes.

Studies	Data	Sample	Methodologies
Abbes (2012)	35 Islamic indices (MSCI) vs their conventional indices	2002-2012 (M)	Sharpe ratio
Ho et al. (2014)	12 Islamic indices their conventional indices	2000-2011 (M)	Sharpe ratio, Jensen's alpha, Treyner ratio
Jawadi et al. (2014)	3 DJ Islamic indexes vs their conventional indices	2000-2011 (M)	Sharpe ratio, Jensen's alpha, Treyner ratio, Omega ratio, Roy ratio Black-Treyner ratio
Ashraf and Mohammad (2014)	12 global and regional Islamic indices (MSCI, DJ and S&P) vs their conventional indices	2002-2012 (M)	LSTAR model
Al-Khazali et al. (2014)	9 DJ Islamic indexes vs their conventional indexes	1996-2012 (M)	stochastic dominance analysis

Notes: D: daily; M: monthly. KLSI: Kuala Lumpur Shariah index; KLCI: Kuala Lumpur Composite index; FTSEGII: FTSE Global Islamic index; FTSEWI: FTSE All-World index; DJIM: DJ Islamic Market index; DJWI: DJ World index; FTSE4G: FTSE For Good index; DJSRWI: DJ Socially Responsible World index.

(2014). Here we use the semi-parametric procedure to detect additive outliers proposed by Laurent, Lecourt and Palm (LLP) (2014).<sup>11</sup> Their test is similar to the non-parametric tests for jumps proposed by Lee and Mykland (2008) and Andersen, Bollerslev, and Dobrev (2007) for daily data. This method allows us to examine the large shocks that affected the returns of Islamic and conventional indexes.

Consider the returns series  $r_t$ , which is defined by  $r_t = \log P_t - \log P_{t-1}$ , where  $P_t$  is the observed price at time  $t$ , and consider the ARMA( $p,q$ )-GARCH(1,1) model

$$\begin{aligned}
 \phi(L)(r_t - \mu) &= \theta(L)\varepsilon_t \quad \text{or} \quad r_t = \mu_t + \varepsilon_t, & (1) \\
 \varepsilon_t &= z_t \sqrt{\sigma_t^2}, \\
 \varepsilon_t &\sim N(0, \sqrt{\sigma_t^2}), \quad z_t \sim i.i.d.N(0, 1), \\
 \sigma_t^2 &= \omega + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2
 \end{aligned}$$

where  $L$  is the lag operator,  $\phi(L) = 1 - \sum_{i=1}^p \phi_i L^i$  and  $\theta(L) = 1 - \sum_{i=1}^q \theta_i L^i$  are polynomials of orders

<sup>11</sup>Laurent et al. (2014) critic the sequential test for outliers proposed by Franses and Ghijssels (1999) in two ways: (i) the critical values have to be simulated and depend on some unknown parameter values of the GARCH model and therefore the size of the test can not be controlled; and (ii) the test suffers from the so-called outlier masking problem because it is based on a Quasi-Maximum Likelihood estimate of the GARCH model, which is known to be non-robust to additive outliers (Carnero et al., 2007, 2012). They show that their test does not suffer from those drawbacks.

$p$  and  $q$ , respectively, such that  $\mu_t = \mu + \sum_{i=1}^{\infty} \lambda_i \varepsilon_{t-i}$  is the conditional mean of  $r_t$ , where  $\lambda_i$ 's are the coefficients of  $\lambda(L) = \phi^{-1}(L)\theta(L) = 1 + \sum_{i=1}^{\infty} \lambda_i L^i$ , and  $\sigma_t^2$  is the conditional variance of  $r_t$ .

Consider the return series with an independent additive outlier component  $a_t I_t$ , with outlier size  $a_t$

$$r_t^* = r_t + a_t I_t \quad (2)$$

where  $r_t^*$  denotes observed financial returns and  $I_t$  is generated by some outlier process such as a Poisson process. The model for  $r_t^*$  has the properties that an outlier  $a_t I_t$  will not affect  $\sigma_{t+1}^2$  (the conditional variance of  $r_{t+1}$ ), and it allows for non-Gaussian fat-tailed conditional distributions of  $r_t^*$ .

Let us denote  $\tilde{\mu}_t$  and  $\tilde{\sigma}_t$  estimates of  $\mu_t$  and  $\sigma_t^2$  in equations (1) and (2) that are robust to the potential presence of the additive outliers  $a_t I_t$  (i.e. estimated on  $r_t^*$  and not on  $r_t$ ). The robust estimations of  $\tilde{\mu}_t$  and  $\tilde{\sigma}_t$  are based on the bounded innovation propagation (BIP)-ARMA proposed by Muler, Peña and Yohai (2009) and the BIP-GARCH(1,1) proposed by Muler and Yohai (2008), respectively. The BIP-ARMA and BIP-GARCH(1,1) are defined as

$$\tilde{\mu}_t = \mu + \sum_{i=1}^{\infty} \lambda_i \omega^{MPY}(\tilde{J}_{t-i}) \quad (3)$$

$$\tilde{\sigma}_t^2 = \omega + \alpha_1 \tilde{\sigma}_{t-1}^2 c_{\delta} \omega_{k_{\delta}}^{MPY}(\tilde{J}_{t-1})^2 + \beta_1 \tilde{\sigma}_{t-1}^2, \quad (4)$$

where  $\omega_{k_{\delta}}^{MPY}(\cdot)$  is the weight function, and  $c_{\delta}$  a factor ensuring the conditional expectation of the weighted squared unexpected shocks to be the conditional variance of  $r_t$  in absence of jumps (Boudt et al., 2013). Laurent et al. (2014) extend the BIP-GARCH to the BIP-GJR(1,1), i.e. the robust version of the GJR(1,1) model of Glosten, Jagannathan, and Runkle (1993) that accounts for the so-called leverage effect, i.e.

$$\tilde{\sigma}_t^2 = \omega + \alpha_1 \tilde{\sigma}_{t-1}^2 c_{\delta} \omega_{k_{\delta}}^{MPY}(\tilde{J}_{t-1})^2 + \gamma_1 D_{t-1} \tilde{\sigma}_{t-1}^2 c_{\delta} \omega_{k_{\delta}}^{MPY}(\tilde{J}_{t-1})^2 + \beta_1 \tilde{\sigma}_{t-1}^2, \quad (5)$$

where  $D_{t-1} = 1$  if  $\tilde{J}_{t-1} < 0$ , and 0 otherwise.

Consider the standardized return on day  $t$

$$\tilde{J}_t = \frac{r_t^* - \tilde{\mu}_t}{\tilde{\sigma}_t} \quad (6)$$

The outliers detection rule is as follow

$$\tilde{I}_t = I(|\tilde{J}_t| > k) \quad (7)$$

where  $I(\cdot)$  is the indicator function, with  $\tilde{I}_t = 1$  when an outlier is detected at observation  $t$  and 0 otherwise, and  $k$  is a suitable critical value. The critical values are defined by

$$k = g_{T,\lambda} = -\log(-\log(1-\lambda))b_T + c_T, \quad (8)$$

with  $b_T = 1/\sqrt{2\log T}$ , and  $c_T = (2\log T)^{1/2} - [\log \pi + \log(\log T)]/[2(2\log T)^{1/2}]$ . They show that their test do not suffer from size distortions irrespectively of the parameter values of the GJR model from Monte Carlo simulations. Following Laurent et al. (2014) we set  $\lambda = 0.5$ . Given  $\tilde{I}_t$ , detected outliers can be filtered out from  $r_t^*$  as follows:  $\tilde{r}_t = r_t^* - (r_t^* - \tilde{\mu}_t)\tilde{I}_t$ .<sup>12</sup>

### 3.2 Sudden change detection

The most popular statistical methods specifically designed to detect breaks in volatility are CUSUM-type tests. As underlined by Rodrigues and Rubia (2011), the ability of the CUSUM tests to identify structural changes depends of the underlying assumptions. Financial data display a time varying volatility patterns, well known as volatility clustering. Andreou and Ghysels (2002) illustrate the pervasive effect persistent volatility on CUSUM-type tests experimentally. Their results indicate that the Kokoszka and Leipus (2000) test has a good properties in presence of conditional heteroscedasticity. Sansó et al. (2004) propose a more general test that one of Kokoszka and Leipus (2000) based on the iterative cumulative sum of squares (ICSS) algorithm developed by Inclán and Tiao (1994).

Let  $e_{i,t} = 100 \times \log(P_{i,t}/P_{i,t-1})$ , where  $P_{i,t}$  is the price of the index  $i$  at the time  $t$ , so that  $e_t$  is the percent return of the index  $i$  from period  $t-1$  to  $t$ .  $\{e_t\}$  is then assumed to be a series of independent observations from a normal distribution with zero mean and unconditional variance  $\sigma_t^2$  for  $t = 1, \dots, T$ . Assume that the variance within each interval is denoted by  $\sigma_j^2$ ,  $j = 0, 1, \dots, N_T$ , where  $N_T$  is the total number of variance changes and  $1 < \kappa_1 < \kappa_2 < \dots < \kappa_{N_T} < T$  are the set of breakpoints. Then the variances over the  $N_T$  intervals are defined as

$$\sigma_t^2 = \begin{cases} \sigma_0^2, & 1 < t < \kappa_1 \\ \sigma_1^2, & \kappa_1 < t < \kappa_2 \\ \dots & \\ \sigma_{N_T}^2, & \kappa_{N_T} < t < T \end{cases}$$

The cumulative sum of squares is used to estimate the number of variance changes and to detect the point in time of each variance shift. The cumulative sum of the squared observations from the beginning

---

<sup>12</sup>Note that the LLP test for additive jumps is based on the assumption that on days without jumps, returns are conditionally Gaussian. It is thus important to use a test for the conditional normality of the filtered returns in addition to their test for jumps. Indeed, a rejection of this assumption might invalidate our postulated DGP and therefore our test. We checked the normality with the Jarque-Bera test.

of the series to the  $k$ th point in time is expressed as  $C_k = \sum_{t=1}^k e_t^2$  for  $k = 1, \dots, T$ . In order to test the null hypothesis of constant unconditional variance, the Inclán–Tiao statistic is given by:

$$IT = \sup_k |(T/2)^{0.5} D_k| \quad (9)$$

where  $D_k = \left(\frac{C_k}{C_T}\right) - \left(\frac{k}{T}\right)$ , with  $C_T$  is the sum of the squared residuals from the whole sample period. The value of  $k$  that maximizes  $|(T/2)^{0.5} D_k|$  is the estimate of the break date. The ICSS algorithm systematically looks for breakpoints along the sample. If there are no variance shifts over the whole sample period,  $D_k$  will oscillate around zero. Otherwise, if there are one or more variance shifts,  $D_k$  will departure from zero. The asymptotic distribution of IT is given by  $\sup_r |W^*(r)|$ , where  $W^*(r) = W(r) - rW(1)$  is a Brownian bridge and  $W(r)$  is standard Brownian motion. Finite-sample critical values can be generated by simulation.

The IT statistic is designed for i.i.d. processes, which is a very strong assumption for financial data, in which there is evidence of conditional heteroskedasticity. Sansó et al. (2004) showed that the size distortions are important for heteroskedastic conditional variance processes from Monte carlo simulations. Their results thus invalidate in practice the use of this test for financial time series. To overcome this problem, Sansó et al. (2004) proposed a new test that explicitly consider the fourth moment properties of the disturbances and the conditional heteroskedasticity.<sup>13</sup> They suggested a non-parametric adjustment to the IT statistic that allows  $e_t$  to obey a wide class of dependent processes under the null hypothesis. As suggested by Sansó et al. (2004), we use a non-parametric adjustment based on the Bartlett kernel, and the adjusted statistic<sup>14</sup> is given by:

$$AIT = \sup_k |T^{-0.5} G_k| \quad (10)$$

where  $G_k = \hat{\lambda}^{-0.5} \left[ C_k - \left(\frac{k}{T}\right) C_T \right]$ ,  $\hat{\lambda} = \hat{\gamma}_0 + 2 \sum_{l=1}^m [1 - l(m+1)^{-1}] \hat{\gamma}_l$ ,  $\hat{\gamma}_l = T^{-1} \sum_{t=l+1}^T (e_t^2 - \hat{\sigma}^2)(e_{t-l}^2 - \hat{\sigma}^2)$ ,  $\hat{\sigma}^2 = T^{-1} C_T$ , and the lag truncation parameter  $m$  is selected using the procedure in Newey and West (1994). Under general conditions, the asymptotic distribution of AIT is also given by  $\sup_r |W^*(r)|$ , and finite-sample critical values can be generated by simulation.

Rodrigues and Rubia (2011) discuss the effects that sample contamination has on the asymptotic properties of CUSUM-type tests for detecting change points in variance and characterize the finite

---

<sup>13</sup>Bacmann and Dubois (2002) show that one way to circumvent this problem is by filtering the return series by a GARCH (1,1) model, and applying the ICSS algorithm developed by Inclán and Tiao (1994) to the standardized residuals obtained from the estimation. Fernandez (2006) propose an alternative approach to testing for variance homogeneity based on wavelet analysis.

<sup>14</sup>This adjusted statistic is equivalent to the non-parametric test proposed by Kokoszka and Leipus (2000).

sample behavior by means of Monte Carlo simulations. They focus on additive outliers, that is exogenous changes that directly affect the series, which prove able to generate large size distortions in these tests. The authors show that the Sansó et al. (2004) test exhibits low power and tends to find few or no breaks at all. Therefore, as suggested by Rodrigues and Rubia (2011), we attempt to identify the variance changes in the stock market indexes from the outlier-adjusted data. Further, Inclán and Tiao (1994) advised that “*it is advisable to complement the search for variance changes with a procedure for outlier detection*”.

## 4 Empirical implementation and results

### 4.1 Description of the data

In February 1999, New York-based Dow Jones was the first to launch Shari’ah-compliant indexes - i.e. based on a subset of investable equities that are compatible with the Islamic finance principles - in respond to the increasing demand for ethical investments from the Muslim community and other socially responsible investors.<sup>15</sup> Broadly speaking, Islamic indexes track the performance of a subset of eligible stocks that are already included in the corresponding global indexes. To become eligible for inclusion in the Islamic index, a company has to satisfy two main screening criteria (see Dow Jones, 2009):

- The *industry* screen, which attempts to remove any companies having primary business activities that are not compatible with the principles of Islamic finance (e.g. alcohol; pork-related products; conventional financial services; entertainment; tobacco; weapons and defense);
- The *financial ratios* screen, which is intended to remove companies based on their levels of leverage or interest income; all of the following financial ratios must not exceed 33% in order for a company to be included in the index: (i) the debt ratio; (ii) the ratio of interest income to total revenue; (iii) the ratio of accounts receivables to the market value of total assets.

In addition, the composition of the indexes is reviewed on a quarterly basis and the index’s weighting scheme follows a free-floating market capitalization, as well as on an ongoing basis to take into account extraordinary events, such us delisting activities, bankruptcies, and mergers. When a new issue is added to the Dow Jones indexes, it is also evaluated according to the Dow Jones Islamic (DJI) indexes criteria to determine whether it will be included in the DJI indexes. All revisions are supervised by an indepen-

---

<sup>15</sup>London-based FTSE, as well as Standard and Poor’s, also launched their specialized Islamic market indexes in 1999, and several years afterwards, respectively.

dent Shari'ah board composed of Islamic scholars.

For the purpose of our analysis, we consider several major (conventional and Islamic) daily stock market indexes spanning January 1<sup>st</sup>, 1996 to March 15<sup>th</sup>, 2013 (4,489 observations) in different jurisdictions: Dow Jones (Islamic) Asian, Dow Jones (Islamic) Canada, Dow Jones (Islamic) Japan, Dow Jones (Islamic) United Kingdom, Dow Jones (Islamic) United States, Dow Jones (Islamic) World. The data are obtained from Thomson Financial Datastream and the Dow Jones company database. The daily returns are computed as the natural logarithmic first difference of the daily closing prices. The logarithmic stock returns are multiplied by 100 to avoid convergence problems.

Table ?? gives the number of stocks used in the calculation of each index, as well as basic descriptive statistics for the original and outlier-adjusted return series. The number of constituents included in the conventional indexes varies from 261 (the DJ UK index) to 6,571 firms (the DJ World index). The application of the filtering criteria reduces the number of stocks included in the Islamic indexes by 60-70% (compared with the broad universe of investable stocks included in the conventional indexes). For the original series, the Islamic market indexes display higher mean returns than the conventional market indexes but they are also slightly more volatile. All the returns are highly non-normal, i.e. showing evidence of negative excess skewness and excess kurtosis. All series are leptokurtic (i.e., fat-tailed distribution) and thus the variance of the index prices is principally due to infrequent but extreme deviations. The Lagrange Multiplier test for the presence of the ARCH effect indicates clearly that the prices show strong conditional heteroscedasticity, which is a common feature of financial data. In other words, there are quiet periods with small price changes and turbulent periods with large oscillations. The outlier-adjusted returns also exhibit excess skewness, excess kurtosis and conditional heteroscedasticity, although the excess kurtosis decreases dramatically, except for the DJ Japan and DJ Islamic UK returns that do not display excess skewness. As shown by Carnero et al. (2001) and Charles and Darné (2005), this result show that outliers may cause significant skewness.

## **4.2 Outliers in daily stock market index returns**

Tables ?? and ?? give the identified outliers for all the indexes in chronological order. In addition, we also associate the date corresponding to each outlier to a specific (economic, political or financial) event that occurred near that date. As expected, outliers have been detected in all the series, and most of them are due to the 2008 financial crisis. Note that the conventional and Islamic indexes display the same



number of outliers, and in the vast majority of cases the corresponding dates are also the same. Over the entire sample period covering 14 years of daily data (January 1<sup>st</sup>, 1996 – March 15<sup>th</sup>, 2013), we show that Islamic and conventional indexes are affected by the same events in most cases<sup>16</sup>, especially the period of the Global Financial Crisis (GFC) of 2007-2008 and its aftermath as well as events associated with the 1997 Asian financial crisis, the Russian crisis of summer 1998, the terrorist attacks in the US on September 11th, 2001, the Japanese earthquake-tsunami in March 2011, as well as other macroeconomic news, such as unanticipated increases in interest rates, consumer price indexes, commodities prices or unemployment rates.<sup>17</sup>

### 4.3 Risk measures

Value-at-risk (VaR) is the most prominent of a set of risk measurement tools developed in response to a series of huge, widely publicized losses at large financial firms in the 80's. From a statistical point of view, VaR is a quantile of the profit-loss distribution, which can be derived for any specific level of significance ( $\alpha$ ) and time horizon. VaR quantifies the potential loss for a portfolio of assets ( $r_t$ ) under normal market condition over a given period of time horizon  $h$  with a certain confidence level ( $1 - \alpha$ ), at time  $t$  conditionally on available information

$$\Omega_{t-1} : P\{r_t \geq VaR_{t,h}(\alpha) | \Omega_{t-1}\}$$

Financial managers estimate the quantile of the left lower-sided tail, as a representation of worst losses for a given ( $\alpha$ ). Providing an accurate estimate of VaR is crucial. If the underlying risk is not properly estimated, this may lead to a suboptimal capital allocation with consequences on the profitability or the financial stability of the institutions and if risk is overestimated, then it may further lead to unnecessary extra capital requirements. The parametric methods for estimating VaR assume one particular distribution for the data series. The Gaussian approach implies that the returns follow a Gaussian distribution. As it is well known that the standard deviation of the returns change over time, we have to use models that explicitly allow the standard deviation to change over time that provide better forecasts of variance and by extension better measures of VaR (Engle, 2011).

We estimate in-sample VaR under 95% confidence levels based on a GJR-GARCH model<sup>18</sup> for the

---

<sup>16</sup>We performed an extensive search on each date using Dow Jones Factiva (Dow Jones and Reuters newswires; key newspapers; and other sources) in order to match each of the identified outliers to one (or several) significant event(s) that occurred on (or near) that date. It is beyond the scope of the paper to discuss at length all these events.

<sup>17</sup>This is consistent with the result reported by Flannery and Protopapadakis (2002), according to which some macroeconomic news announcements have a significant impact on the stock market returns.

<sup>18</sup>We do not search the best volatility models for computing VaR, but this point will be examined in future research.

conventional and Islamic indexes, and we use mean VaR ( $\overline{VaR}$ ) for comparison. However, the VaR is not considered as a coherent measure of risk in the sense of Artzner et al. (1999).<sup>19</sup> Expected shortfall is a coherent measure of risk and it is defined as the expected value of the losses conditional on the loss being larger than the VaR. The expected shortfall is defined as:

$$ESF_t = E(|L_t| > |VaR_t|)$$

where  $L_t$  is the expected value of loss if a  $VaR_t$  violation occurs. Table ?? displays the measures of  $\overline{VaR}$  and ESF as well as simple risk measures, standard deviation and semivariance (downside risk). The results show that the risk measures are higher for the Islamic indexes than for their conventional counterparts, suggesting that the Islamic indexes are more risky. This result is opposed to that of Al-Zoubi and Maghyereh (2007) who found that the risk level of Islamic index is lower than their conventional counterpart.

We also compute the risk measures from adjusted-outlier data because the presence of outliers may have undesirable effects on the estimates of the parameters of the equation governing volatility dynamics (see, e.g., Van Dijk et al., 1999; Carnero et al., 2007; Charles, 2008) and the measures of risk (see, e.g., Iqbal and Mukherjee, 2012; Charles and Darné, 2014; Grané and Veiga, 2014). We obtain the same results than from the original data, but with a lower level of risk.

#### 4.4 Performance measures

The standard performance measures developed in the literature consist of investigating the relationship between the expected returns and risk associated with investment in risky financial assets. Several tools have been introduced to evaluate stock market index performance and these often differ depending on the type of risk measure under consideration. We apply different standard performance measures, namely the Sharpe ratio, Treynor ratio, Jensen's alpha and Black-Treynor ratio, and two improved Sharpe ratios based on others proxies of risk (semivariance and VaR) to conventional and Islamic indexes.

The first performance measure is the Sharpe ratio ( $SR$ ), also often referred to as "Reward to Variability", which indicates if an investment's high return is a result of excessive risk. It measures the performance of an index by dividing the amount of excess return (risk premium) to total risk, measured by standard deviation. The higher the  $SR$  is consistent with a higher probability that the index return

---

<sup>19</sup>In the properties a coherent measure functional must satisfy on an appropriate probabilistic space, the sub-additivity property does not hold for all cases. Specific portfolios can be constructed where the risk of a portfolio with two assets can be greater than the sum of the individual risks therefore, violating sub-additivity and in general the diversification principle (Scaillet, 2000).

exceed the risk-free return. If the  $SR$  is negative (resp. positive), the index  $i$  underperforms (resp. outperforms) the referential given by the risk-free asset.

$$SR = \frac{(R_{i,t} - R_{f,t})}{\sigma(R_{i,t})}$$

where  $R_{i,t}$  denotes the stock return of the index  $i$ ,  $R_{f,t}$  refers to the risk-free return, and  $\sigma(R_{i,t})$  the standard deviation of the returns of the index  $i$ .

The second performance measure is the Treynor ratio ( $TR$ ) which measures the index performance for its given level of market risk (CAPM). The  $TR$  is similar to the  $SR$  except that it substitutes total risk by systematic risk and therefore uses beta instead of standard deviation as a proxy for risk.

$$TR = \frac{(R_{i,t} - R_{f,t})}{\beta_i}$$

where  $\beta_i = cov(R_{i,t}, R_{M,t}) / \sigma^2(R_{M,t})$ , with  $R_{M,t}$  the market benchmark return. We use the MSCI AC World index as the market benchmark.

The third risk-adjusted performance measure is the adjusted Jensen's alpha ( $\alpha$ ), also based on systematic risk. It measures the performance excess of a fund or portfolio in relation to the CAPM performance. A positive  $\alpha$  means that the index achieves excess return relative to the market, negative  $\alpha$  means underperformance.

$$\alpha_i = (R_{i,t} - R_{f,t}) - \beta_i(R_{M,t} - R_{f,t})$$

The fourth performance measure is the Black-Treynor ratio ( $BTR$ ), which can be defined as the alpha-beta ratio of a portfolio.

$$BTR = \frac{\alpha_i}{\beta_i}$$

The fifth risk-adjusted performance measure is the Reward-to-VaR ratio ( $RVaR$ ) proposed by Dowd (2002). It is similar to the  $SR$  and  $TR$  but it uses another proxy for risk, the VaR.

$$RVaR = \frac{(R_{i,t} - R_{f,t})}{VaR_i}$$

where  $VaR_i$  represents the VaR measure of the the index  $i$ .

The last performance measure is the Sortino ratio ( $SOR$ ) which is a modification of the  $SR$  but penalizes only the returns falling below a user-specified target or required rate of return, while the  $SR$  penalizes both upside and downside volatility equally.

$$SOR = \frac{(R_{i,t} - R_{f,t})}{SV(R_{i,t})}$$

where  $SV(R_{i,t})$  is the semivariance (downside risk) of the returns of the index  $i$ .

Table ?? displays the risk-adjusted performance measures for the Islamic and conventional indexes. The results show that in most cases the Islamic indexes seem to exhibit higher risk-adjusted performance than their conventional counterpart, whatever the measures, except for the World index for what the results are mixed. This finding confirms those of Ashraf and Mohammad (2014) and Jawadi et al. (2014). Note that the performance measures based on outlier-adjusted series give the same results.

## 5 Risk and performance measures in volatility changes

The time periods of a shift in volatility as detected by the modified ICSS algorithm from the outlier-adjusted data are given in Table ?. Variance breaks are identified in all conventional and Islamic indexes, ranging from two to six shifts. The conventional and Islamic indexes displayed the same number of variance changes with the same dates, except for the emerging market indexes for which the dates are slightly different. For example, three sudden changes have been detected both in the DJ Asia and DJ Islamic Asia series, with the following dates: 11/27/1996, 10/23/2002 and 07/26/2007. This result tends to confirm that the Islamic indexes have not been more or less affected by variance changes than the conventional indexes.

As noted by Malik and Hassan (2004), some major events correlate with the regime shifts as identified by the ICSS algorithm, and some events are often marked by the peaks and lows of the market. We believe that these events are contributing factors. However, markets may very well anticipate some events in advance and sometimes respond with a time lag, so we do not expect these events to correlate to changes in sudden variance on any specific day. In this paper, we only suggest that these events may be a contributing factor in the sudden change and make no attempt to show definitely the causes of the sudden changes. These structural breaks in volatility allow to define low volatility periods and high volatility periods.

We compare the measures of risk and risk-adjusted performance between both indexes from the different sub-periods, based on adjusted-outlier data (Tables ??–??).<sup>20</sup> The results show that most of the Islamic indexes have higher level of risk than the conventional indexes, whatever the sub-periods. This result is opposed to that of Hussein (2004), Dharani and Natarajan (2011) and Ho et al. (2014) who found that the Islamic indexes are less risky than the conventional index during crisis periods.<sup>21</sup>

---

<sup>20</sup>We also computed the risk measures from original data and obtained the same results than from the adjusted-outlier data. The results are available upon request.

<sup>21</sup>Their measure of risk is only based on the systematic risk ( $\beta$ ).

For example, the DJ Canada index exhibits a standard deviation and a mean VaR of 1.52 and -0.0273, respectively, over the 07/31/1998–04/23/2001 period, whereas those of the DJ Islamic Canada index are of 2.51 and -0.0347, respectively. Further, for all the indexes the more risky sub-period is associated with the GFC, due to a very high level of uncertainty. Consequently, this finding means that the Islamic indexes are riskier than the non-Islamic indexes. We also find that in most subperiods the Islamic indexes either outperform the non-Islamic indexes or there is no significant difference in performance between both indexes, from almost all the risk-adjusted measures.

## 6 Discussion

Alternative plausible explanations have been proposed in the literature to explain the difference of risk between the Islamic and conventional indexes. First, the relative under-diversification of the Islamic indexes due to filtering criteria that remove a large number of *Shari'ah* non-compliant firms.<sup>22</sup> Indeed, the screening reduces the number of stocks included in the DJ Islamic indexes by 60-70% (Table ??). As a result, *Shari'ah* compliant firms becomes less diversified and being concentrated in some specific sectors, which further results in more volatile returns (Hussein and Omran, 2005). Therefore, lower leverage and less diversification are the main distinctive features of Islamic indices.<sup>23</sup> By looking at the constituent list of DJ Islamic indexes across different sectors (Table ??), *Shari'ah* filtering has reduced the composition of firms listed in these indexes. The consequence of less diversification in Islamic indexes is to lead to higher concentration risk in some sectors. Most of DJ Islamic indexes are driven by investing in few sectors, such as basic material, industrial and technology firms.<sup>24</sup> We further can

---

<sup>22</sup>There is limit to diversification and the exclusion of *Shari'ah* non-compliant firms may not necessarily mean that Islamic stock indexes are under-diversified. The different combination of stocks also matters. It is possible that the tobacco and alcohol industries have lesser volatility and higher returns in view of the relatively stable consumption behavior. It is also possible that conventional financial sector stocks are more volatile and respond more sensitively to shocks. We thank the referee for this comment.

<sup>23</sup>However, Girard and Hassan (2008) and Kok et al. (2009) suggested to use Islamic and conventional indices as asset classes to have more diversification benefits. Kok et al. (2009) and Seng et al. (2009) exhibited the existence of diversification opportunities by grouping Islamic with conventional and socially responsible indices on the UK market. Majid and Kassim (2010) find that the investors can gain benefits by diversifying in the Islamic stock markets across economic groupings such as that in the developed and developing countries, but limited benefits within the same economic groupings.

<sup>24</sup>Dewandaru et al. (2015) find that DJ Islamic indexes are concentrated in health care, industrial goods, oil and gas, retail and technology. Girard and Hassan (2008) find that FTSE Islamic indices are growth oriented and conventional indices are relatively more value focused, and suggest that one reason behind the high proportion of growth stocks may come from the exclusion of value sectors with higher environmental risks, such as chemical, energy, and basic industries.

observe a few exceptions for some countries in Islamic indexes with oil and gas sector for Canada and the UK, and health care sector for the US. The DJ conventional indexes are more concentrated in industrial, consumer services and financial sectors. The DJ and DJI of the Canada indexes are strongly focused on oil and gas and basic material firms that can explain their higher level in standard deviation and semi-variance compared to other jurisdictions, implying lower Sharpe and Sortino ratios based on these two proxies of risk. Both Islamic and conventional indexes for Asia, Japan and World display higher measures of risk-adjusted performance based on total risk (Sharpe and Sortino ratios) than other jurisdictions (due to low total risks) which could be explained by the fact that they have a higher proportion of consumer good firms than other jurisdictions. The DJ and DJI US indexes have a higher proportion of health care and technology firms than other markets that could explain their higher measures of risk-adjusted performance based on systematic risk (Treydor ratio, Jensen's alpha and Black-Treydor ratio).<sup>25</sup>

Second, the systematic exclusion of the largest firms from the broad universe of investable equities included in Islamic indexes due to the financial ratios screen, which implies that the remaining *Shari'ah* compliant firms are smaller (Hussein and Omran, 2005; Girard and Hassan, 2008; Sati et al., 2014). Hussein and Omran (2005) and Girard and Hassan (2008) find that Islamic indices are small-cap oriented and conventional indices are relatively more mid-cap focused. Table ?? also reports the sub-indexes based on size, large-, mid- and small-caps, for the conventional and Islamic indexes. In contradiction with the previous studies, all conventional and Islamic indexes are small-cap oriented, except for DJ Islamic Asia and Japan indexes which are more mid-cap oriented. Further, the *Shari'ah* compliant screens slightly modify the proportion of firm sizes (increasing or decreasing). Therefore, it seems that a potential size bias due to the *Shari'ah* filtering criteria, which tend to exclude the largest, more stable, constituents from the broad universe of investable stocks, cannot be a relevant explanation of the higher risk of Islamic indexes.

Langbein and Posner (1980) argue that ethical investment may involve higher risk but should not yield significantly worse returns since ethical investors do not invest in clearly unprofitable stock.

## 7 Conclusion

The present paper contributes to the literature on the impact of the *Shari'ah* filtering criteria on the risk of Dow Jones Islamic indexes relative to their conventional counterparts. We first analyzed whether

---

<sup>25</sup>The higher Treynor ratio, Jensen's alpha and Black-Treydor ratio for the DJ and DJI US indexes compared to other markets are due to their lower beta (systematic risk). The estimated betas are available upon request.

the Islamic indexes were more affected by extreme events than the conventional indexes, using outlier detection procedure and several major Dow Jones Market indexes. Then, we examined whether the Islamic indexes were more risky than the conventional indexes using different risk measures. We also analyzed the performance of both indexes from various risk-adjusted performance measures. The sample period (1996-2013) was further divided into low volatility period and high volatility period based on the detection of structural breaks in the volatility.

We showed that Islamic and conventional indexes were affected by the same events in most cases which can bias the estimation of the risk, especially the period of the Global Financial Crisis (GFC) of 2007-2008 and its aftermath which was characterized by a very high level of volatility. Overall, the Islamic indexes seems to be more risky than their conventional counterparts as well as exhibit a higher performance on the full period. The sample period is further divided into low volatility period and high volatility period based on the detection of structural s. We show that most of the Islamic indexes have higher level of risk than the conventional indexes, whatever the sub-periods. Consequently, this finding means that the Islamic indexes are riskier than the non-Islamic indexes. We also find that in most cases the Islamic indexes either outperform the non-Islamic indexes or there is no significant difference in performance between both indexes. These findings can be explained as a consequence of less diversification in Islamic indexes, leading to higher concentration risk in some sectors, such as basic material, industrial and technology firms.

Further, we also find some differences between the jurisdictions. The DJ and DJI of the Canada indexes exhibited lower Sharpe and Sortino ratios, due to higher level in standard deviation and semi-variance, compared to other jurisdictions. Both Islamic and conventional indexes for Asia, Japan and World display higher measures of risk-adjusted performance based on total risk, due to low total risks, than other jurisdictions. The DJ and DJI US indexes have higher measures of risk-adjusted performance based on systematic risk (Treynor ratio, Jensen's alpha and Black-Treynor ratio) than other markets. These differences of risk and performance seem to be explained by the fact that these markets are oriented in some sectors.

These explanations on the fact that Islamic indexes and markets are sector oriented will be the subject of further research.

## References

- [1] Abbes M.B., (2012). Risk and return of Islamic and conventional indices. *International Journal of Euro-Mediterranean Studies*, 5, 1-23.
- [2] Abdelsalam O., Duygun M., Matallín-Sáez J.C. and Tortosa-Ausina E., (2014). Do ethics imply persistence? The case of Islamic and socially responsible funds. *Journal of Banking and Finance*, 40, 182-194.
- [3] Ahmad Z. and Ibrahim H., 2002. A study of performance of the KLSE Shariah index. *Malaysian Management Journal*, 6, 25-34.
- [4] Ahmed A., 2010. Global financial crisis: an Islamic finance perspective. *International Journal of Islamic and Middle Eastern Finance and Management*, 3, 306-320.
- [5] Ajmi A.N., Hammoudeh S., Nguyen D.K. and Sarafrazi S., 2014. How strong are the causal relationships between Islamic stock markets and conventional financial systems? Evidence from linear and nonlinear tests. *Journal of International Financial Markets, Institutions and Money*, 28, 213-227.
- [6] Alasrag, H., 2010. Global Financial Crisis and Islamic finance. Working Paper.
- [7] Albaity M. and Ahmad R., 2008. Performance of Shariah and composite indices: Evidence from Bursa Malaysia. *Journal of Accounting and Finance*, 4, 23-43.
- [8] Albaity M. and Ahmad R., 2011a. Return performance and leverage effect in Islamic and socially responsible stock indices: Evidence from Dow Jones (DJ) and Financial Times Stock Exchange (FTSE). *African Journal of Business Management*, 5, 6927-6939.
- [9] Albaity M. and Ahmad R., 2011b. Return performance, leverage effect, and volatility spillover in Islamic stock indices evidence from DJIMI, FTSEGII and KLSI. *Investment Management and Financial Innovations*, 8, 161-171.
- [10] Al-Khazali O., Lean H.H. and Samet A., 2014. Do Islamic stock indexes outperform conventional stock indexes? A stochastic dominance approach. *Pacific-Basin Finance Journal*, 28, 29-46.
- [11] Al-Zoubi H. and Maghyereh A., 2007. The relative risk performance of Islamic finance: a new guide to less risky investments. *International Journal of Theoretical and Applied Finance*, 10, 235-47.



- [12] Andersen T.G., Bollerslev T. and Dobrev D., 2007, Noarbitrage semi-martingale restrictions for continous-time volatility models subject to leverage effects, jumps and i.i.d. noise: Theory and testable distributional implications. *Journal of Econometrics*, 138, 125-180.
- [13] Andreou E. and Ghysels E., 2002. Detecting multiple breaks in financial market volatility dynamics. *Journal of Applied Econometrics*, 17, 579-600.
- [14] Artzner P., Delbaen F., Eber J-M. and Heath D., 1999. Coherent measures of risk. *Mathematical Finance*, 9, 203-228.
- [15] Ashraf D. and Mohammad N., 2014. Matching perception with the reality-Performance of Islamic equity investments. *Pacific-Basin Finance Journal*, 28, 175-189.
- [16] Bacmann J.F. and Dubois M., 2002. Volatility in emerging stock markets revisited. European Financial Management Association, London Meeting, 26-29 June 2002.
- [17] Bali R. and Guirguis H., 2007. Extreme observations and non-normality in ARCH and GARCH. *International Review of Economics and Finance*, 16, 332-346.
- [18] Balke N.S. and Fomby T.B., 1994. Large shocks, small shocks, and economic flucutuations: Outliers in macroeconomic time series. *Journal of Applied Econometrics*, 9, 181-200.
- [19] Black F., 1976. Studies of stock price volatility change. *Proceedings of the 1976 Meetings of the Business and Economics Statistics Section, American Statistical Association*, 177-181.
- [20] Bollerslev T., Chou R.Y. and Kroner K.F., 1992. ARCH modeling in finance: A review of the theory and empirical evidence. *Journal of Econometrics*, 52, 5-59.
- [21] Boudt K., Danielsson J. and Laurent S., 2013. Robust forecasting of dynamic conditional correlation GARCH Models. *International Journal of Forecasting*, 29, 244-257.
- [22] Box G.E.P. and Tiao G.C. (1975). Intervention analysis with applications to economic and environmental problems. *Journal of the American Statistical Association*, 70, 70-79.
- [23] Carnero M.A., Peña D. and Ruiz E., 2001. Outliers and conditional autoregressive heteroskedasticity in time series. *Revista Estadística*, 53, 143-213.
- [24] Carnero M.A., Peña D. and Ruiz E., 2007. Effects of outliers on the identification and estimation of the GARCH models. *Journal of Time Series Analysis*, 28, 471-497.