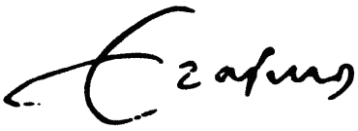
MSc Thesis Behavioural Economics

December 2014



Impaired Auditor's Credibility and its Effect on the Stock Market: Evidence from The Netherlands

By L. (Lars) in 't Veld (333534), Supervised by Prof. Dr. H. (Han) Bleichrodt

Abstract: *This study examines whether impairment of auditor's reputation negatively affects stock returns of the publicly listed auditor's clients. By conducting an event study, the impact of six enforcement initiatives imposed by Dutch authorities are analyzed using the market model on companies listed on the Dutch Euronext Amsterdam. Moreover, the differences in stock returns of the auditor's clients and the non-clients are tested using both parametric and nonparametric tests. Although the results are mixed, the outcomes hint that the auditor's clients perform worse than the non-clients during the days following the event. In case any market reaction is detected, it is observed on the event date directly.*

TABLE OF CONTENTS

Impaired Auditor's Credibility and its Effect on the Stock Market: Evidence from The Netherlands

By L. (Lars) in 't Veld (333534), Supervised by Prof. Dr. H. (Han) Bleichrodt[1]

Erasmus University Rotterdam, Erasmus School of Economics, MSc Economics and Business, Behavioural Economics: Financial Economics

Abstract: *This study examines whether impairment of auditor's reputation negatively affects stock returns of the publicly listed auditor's clients. By conducting an event study, the impact of six enforcement initiatives imposed by Dutch authorities are analyzed using the market model on companies listed on the Dutch Euronext Amsterdam. Moreover, the differences in stock returns of the auditor's clients and the non-clients are tested using both parametric and nonparametric tests. Although the results are mixed, the outcomes hint that the auditor's clients perform worse than the non-clients during the days following the event. In case any market reaction is detected, it is observed on the event date directly.*

I. Introduction

"*An accountant's greatest asset is its reputation for honesty, closely followed by its reputation for careful work.*"[2] However, these specific auditors' assets seem to be impaired significantly since the world was shocked by several international accounting scandals around the year 2000, e.g. Enron, Worldcom and Parmalat (Barton, 2005). Nevertheless, audit failures are not bound by geography nor time. In the period before the millennium change as well as during more recent years, we have seen a vast amount of cases where accounting rules were applied incorrectly, of which some represented fraud cases.

The ongoing occurrence of similar cases suggests that in some events also accounting firms may have certain incentives to reduce their audit efforts and hereby enable client's management to apply accounting rules incorrectly or allow management to stretch these rules. The incentives accounting firms may have to reduce audit efforts could probably be explained by some basic (business) economic principles, varying from reducing audit efforts and hereby costs – and thus increasing profit – to client relationship management rationales. In short, the benefits related to auditor acquiescence in managerial

[1] Prof. dr. Han Bleichrodt works at the Erasmus School of Economics, and is particularly specialized in (health) utility management, decision under uncertainty, intertemporal choice, and social choice theory. I would like to gratefully thank Han for his patience and helpful assistance during the writing process of my thesis.

[2] Quoted from Judge Easterbrook, in the case between Ernst & Young and DiLeo (901 F.2d 624 (7th Cir. 1990)).

misbehavior seem to exceed its (potential) costs in certain cases in the short term (Coffee, 2002). Or, as Levitt (1998) once expressed, "*auditors are participants in a game of nods and winks*."

Mainly due to the aforementioned scandals and their impact on the entire financial system, trust in auditors seems to be impaired.[3] As such, this impairment seems to have a negative impact on the society's perception of these so-called gatekeepers[4]. Positive auditor reputation, however, is a *conditio sine qua non* for the long-run business viability of the audit firm (Callen & Morel, 2002). The auditor eliminates some of the (investor) uncertainty around his clients' financial performance by auditing – i.e. verifying and certifying – their clients' financial statements. However, in case the reliability of the auditor and his audit services is questioned, we assume that the uncertainty around the client's financial performance would increase instead of decrease.[5] Given that most investors are ambiguity averse (Fox & Tversky, 1995), we hypothesize that the impairment of auditor reputation will negatively affect the stock performance of his clients. This study tests this hypothesis and examines whether the impairment of auditor reputation has a negative impact on the stock returns of the auditor's clients. Given that (the impairment of) auditor reputation is not readily observable, a variety of publications of enforcement initiatives imposed by Dutch supervising authorities on audit firms, serve as a proxy for auditor reputation that is impaired. This study considered only six proxies, as auditors are not reprimanded frequently since the supervision on auditors is introduced in 2006.[6] Hence, the current study addresses the following research question: Do the auditor's clients, other than the relevant audited client involved, experience negative abnormal stock returns on the news that their external auditor is penalized by one of the Dutch (financial) authorities?

By conducting an event study, we try to answer the research question defined above. The underlying hypothesis is that investors take into account the (impaired) reputation of the company's external auditor in assessing the robustness of financial statements and, subsequently, in making their investment decisions. In case the underlying hypothesis is confirmed, the stock returns of the auditor's clients would be negatively affected by impaired auditor reputation caused by disciplinary sanctions.

The results are mixed. Although the auditor's clients do not experience a consistent market reaction following the events, the study's results suggest that the auditor's clients perform worse than their control group. In case investors (seem to) react, the market reaction is detected on the event date directly. Hence, the consequences of stressed auditor reputation seem to spill over onto the auditor's clients.

The remainder of this study is organized as follows. Section II surveys the theoretical background and presents the main research hypothesis. Section III describes the research methodology. Section IV provides an analysis of the data. Section V discusses this study's results. Section VI evaluates and concludes.

[3] De Wit *et al.* (2010) addressed in their parliamentary study some critical remarks on the role of the auditor. Barnier, on behalf of the European Commission (2010), launches a public consultation on audit policy. Hence, the role of the auditor within the corporate governance of (listed) companies is politically criticized.

[4] Coffee (2002) refers to the auditor with the term 'gatekeeper' to specify the auditor's main role. Essentially, the gatekeeper is an reputational intermediary who provides verification and certification services to investors – e.g. shareholders – by verifying and certifying its client's financial statements. Since it is assumed that the (theoretically independent) gatekeeper has a lesser incentive to acquiesce in financial statement fraud than its client, the stakeholders regard the gatekeeper's evaluation as more credible.

[5] E.g. publications on inadequate execution of audit principles by auditors suggest that this would increase the uncertainty regarding the credibility of the client's financial statements, since it rises questions whether the audit of the clients' financial statements is credible.

[6] Supervision on auditors is introduced as the Audit Firms Supervision Act (Wta) became effective on October 1, 2006.

II. THEORETICAL BACKGROUND

This section depicts the theoretical background regarding the auditor's role and the consequences of potential impaired auditor reputation. For this purpose subsection II.I discusses the auditor's role – i.e. verifying and certifying financial statements – and the consequences of impaired auditor reputation on the credibility of the auditor's auditing services. Subsection II.II discusses empirical studies on the stock market impact of impaired auditor reputation. The central research hypothesis is presented in subsection II.III.

II.I. A BROAD PERSPECTIVE: AUDITORS AND AMBIGUITY

As aforementioned, Coffee (2002) enlightens that the auditor should – both ideally and theoretically – serve as an independent gatekeeper in order to provide verification and certification services to investors.[7] By doing so, the gatekeeper provides credible assurance on the reliability of the company's financial statements.[8] Since the investor – e.g. the company's shareholder – and the company's management have conflicting interests,[9] the auditor's main role is to approve the company's financial statements and – hereby – to indirectly audit the company's management. Hence, the auditor is the mandatory[10] and the – at least theoretically – independent agent of the investor who should reduce the incentive problems that arise when the investors' and management's interests are not aligned. These situations occur when the company's management does not own (all) shares of the company (Jensen & Meckling, 1976; Watts & Zimmerman, 1983).

The verification and certification services that auditors provide to the client's investors can be defined as credence goods. This suggests that investors can never be fully certain of the exact quality of the auditor's performance. Although the performance may be perceived as credible, whether the auditor really provides high quality service is hardly verifiable. In this regard, auditor reputation serves as primary quality indicator for audit services. A positive – hence, not impaired – auditor reputation can be perceived as quality assurance of providing adequate audit services. In that sense, positive auditor reputation is crucial and seems to be a *conditio sine qua non* for the long-run business viability of the audit firm (Callen & Morel, 2002; Doogar, Sougiannis & Xie, 2003).

According to the existing literature, the auditor's incentives to provide adequate audit services can be distinguished in the insurance and the reputation rationale. The first implies that auditors could be sued by users of the financial statements for damages suffered due to audit failures. The threat of being sued in cases of audit failures direct auditors to deliver adequate audit services. Since bigger audit firms have 'deeper pockets', and thus are more likely to be sued for larger amounts in case of audit failures, these firms have larger incentives to provide high quality audit services. The reputation rationale implies that an auditor should maintain his positive reputation for quality, since a reputation for providing high

[7] In line with Barton (2005), we choose the term 'investors' instead of 'shareholders' in order to refer to all type of investors, e.g. shareholders and creditors.
[8] Note, that although the auditor is appointed and paid by the company that he is to watch, it is assumed that the auditor reputation of providing credible services to investors, is not sacrificed for a single client or fee (Coffee, 2002). This kind of reasoning is called the 'reputation rationale' in literature and is discussed in this section as well.
[9] We note that also other conflicts of interests may exist, for example between the company's debt holder and management (Jensen & Meckling, 1976).
[10] All companies in this study's dataset are obliged to audit their financial statements by a certified auditor, according to Dutch law, i.e. art. 2:393 of the Dutch Civil Code (DCC).

quality audit services is a necessary condition to maintaining their current, and attracting potential new, clients (Weber, Willenborg & Zhang, 2008). DeAngelo (1981) relates the reputation rationale to the risk of losing client-specific quasi-economic rents, since this risk due to an impaired reputation prevents auditors from opportunistic behavior and induces auditors to perform quality audits. In short, auditors are incentivized to provide high quality audit services, since there is a possible threat of being sued and losing attractiveness towards (new) audit clients according to the insurance and reputation rationale respectively.

This subsection's remainder focuses on the auditor's task to reduce (some) investor uncertainty on the client's financial performance by verifying and certifying the clients' financial statements. Since investors cannot readily verify the actual company's financial performance, decisions whether to invest in company's equity are characterized by a certain degree of uncertainty. Company's financial statements serve as performance indicators and – thus – could reduce the investor uncertainty. By auditing the company's financial statements, the auditor's main role is to approve the documents and provide a certain degree of assurance that these documents are free of material misstatements (Doogar *et al.*, 2003).[11] Chaney and Philipich (2002) describe that the audit report does not longer provide an appropriate level of assurance when the news reaches the market that the quality of the auditor's performance is not up to standard. Subsequently, questions arise whether the audited financial statements reflect the actual financial reality and the degree of uncertainty concerning the credibility of the company's financial statements would increase instead of decrease (Weber *et al.*, 2008).

Fox and Tversky (1995) find that most people – investors – dislike uncertainty, in particular if there is a risky alternative with known probabilities available. This ambiguity aversion, i.e. the phenomenon that a risky prospect with known probabilities is preferred to an uncertain prospect with unknown probabilities, seems to vanish in a non-comparative context.[12] In short, Fox and Tversky conclude that ambiguity aversion is driven primarily by a comparison between events and is largely eliminated in the absence of such a comparison. Moreover, most investors respond asymmetrically to economic information: news publications concerning bad news provide a much greater negative effect than positive news publications affect positively (e.g. Soroka, 2006). This phenomena of choice under both risk and uncertainty that – in absolute terms – losses have a larger impact than gains is referred to as loss aversion (Kahneman & Tversky, 1984; Tversky & Kahneman, 1991).

Such comparative context is faced in the case that the reputation of a client's auditor is impaired. As discussed before, the financial performance of companies is less certain. That implies that the chance of a positive return (or negative return) is not being readily observable for the investor is more present. Auditors should eliminate (some) uncertainty by approving – i.e. verifying and certifying – the company's financial statements. However, when the news that the quality of an audit service is no longer considered

[11] Consistent with DeAngelo (1981) and Barton (2005), we note that the auditor cannot entirely ensure that the auditor will uncover and expose material misstatements in the audited financial statements. Hence, audit quality is generally viewed as the probability that the auditor will both uncover and expose such misstatements.

[12] The so-called Ellsberg paradox practically describes ambiguity aversion. Fox & Tversky (1995) shows that a risky prospect with a 50% probability of success (or non-success) is preferred to an uncertain probability of success (or non-success). Neoclassical – rational – economics, however, would suggest that people would be indifferent. In case the comparative context disappears, i.e. the case the subject cannot choose between the risky and uncertain bet, and should have state his willingness to pay for each bet separately, the ambiguity aversion vanishes.

appropriate reaches the market, the uncertainty on the credibility of the company's financial statements will increase. By distinguishing between the particular auditor's clients and other companies, a comparable context is created and the effect of the impaired auditor reputation can be measured. Investors could invest in the equity of either the auditor's clients or his non-clients, whereas the decision to invest in the first would be more uncertain than investing in the latter. Based on the ambiguity averse and loss averse characteristics of investors, investing in the auditor's clients would be less attractive under our hypothesis. We schematically summarize the consequences of auditor reputation impairment in the figure below.

Figure I: Consequences of Reputation Impairment

This figure depicts the relationship between reputation impairment and its consequences for audit clients (treatment group) and non-clients (control group).

II.II. A Specific Perspective: The Impact of Auditor Reputation

Whereas the previous subsection discusses the auditor's position within the corporate governance framework and his task to reduce a certain degree of uncertainty, this subsection specifically focuses on the impact of auditor reputation on the stock performance of the auditor's clients. This subsection discusses some studies that have already assessed the impact of auditor reputation on its clients' stock performance. We note that this subsection only highlights a selection of relevant literature and does not aim to provide a complete and exhausting overview of the findings.

As one of the pioneers, Firth (1990) assesses the empirical impact of various audit reviews by analyzing the impact of quasi-legal reports, issued by the British Department of Trade (DoT)[13]. Hence, the publication of the DoT-reports serve as proxy for impaired auditor reputation. As hypothesized in the present study, it is assumed that the DoT-reports question the quality of the auditor's activities and harm auditor reputation. Subsequently, the credibility of the audited financial statements will impair. Basically, Firth hypothesizes that these DoT-reports lead to a negative market reaction for the shares of the auditor's clients, others than the company involved. Firth's findings suggest that this kind of reputational damage leads to economic losses for both the specific auditor and its clients.[14] For this study, in particular the latter is relevant. The average negative abnormal return, based on the aggregation of all DoT-reports,

[13] The DoT-reports typically describe the history of the criticized company, the (lack of) quality of the company's management and attach responsibilities (Firth, 1990). The study makes use of 16 DoT-reports that have also criticized the auditor's performance. These criticisms are divided in four categories, varying from the lack of independence to an audit of insufficient quality.
[14] Firth (1990) finds a reduction in market share of the auditor following the critical DoT-report as well. These findings suggest that the auditor experiences losses after the publication of the DoT-report.

is significant at a 5% significance level during the week of publication. During that period, the auditors' clients experience an average negative abnormal return of -1.1%. In short, Firth concludes that the publications of the DoT-report lead to a negative market reaction, which is only observable in the week directly following the publication. This confirms the theoretical explanation that investors increase their concerns about the credibility of the clients' financial statements following the publication of DoT-reports.

Callen and Morel (2002) assess the impact of auditor reputation as well. However, they study an alternative context and come to different conclusions than Firth. Callen and Morel's study centralizes around the Enron-Andersen debacle. After analyzing four months of data around news on the Enron scandal, they conclude that the daily abnormal returns are not significantly different from zero (0) after news publication days for both Andersen's clients and the constructed control group, i.e. the Andersen's non-clients. In short, Callen and Morel do not find any market reaction. Moreover, in case any significant negative returns are detected, the abnormal performance of the auditor's clients is not significantly worse than the performance of Andersen's non-clients. An explanation may be that when the events have affected the auditor's clients performance negatively, a so-called spillover effect onto the rest of the market may be detected, for example due to overall deterrence in trust in the financial markets. The existence of such spillover effect, which suggests that failures by one auditor can imperil the entire market for auditing, is confirmed in the literature (Firth, 1990; Doogar *et al.*, 2003). Despite of the conclusion that Andersen's clients do not perform significantly negative after the events, the accumulation of all event days suggests that these clients experience significantly lower cumulative abnormal returns than the control group does. The latter conclusion is mainly due to the market reaction after the events that are directly related to Andersen. These events seem to have a larger (negative) impact on clients' stock returns than events related to the particular client involved, i.e. Enron.

In line with Callen and Morel, Chaney and Philipich (2002) investigate the impact of the Enron-Andersen case and also conclude that events directly related to the specific auditor lead to a statistically negative market reaction. This suggests that, as a reaction on these events, investors downgrade the quality of the audit services and thus the credibility of the audited company's financial statements. Chaney and Philipich centralize January 10, 2002, i.e. the day Andersen (implicitly) admitted guilt, as main event date. The information released on that day in particular seems to cast a disparaging image on Andersen's reputation and – more specifically – its audit procedures. Their results suggest that Andersen's clients indeed experience statistically negative returns following the central event.

Furthermore, these conclusions seem not to be bounded by the sizeable Enron-Andersen affair only. Based on the ComROAD AG-KPMG scandal, Weber *et al.* (2008) draw comparable conclusions. They find that KPMG's clients experience a negative cumulative abnormal return of 3% at a 1% significance level at events pertaining to ComROAD. This suggests that German investors value auditor reputation.

All in all, the discussed literature suggests that the stock market takes auditor reputation – and especially the impairment thereof – into account. Although the results are mixed, evidence for a negative market reaction is found in the UK (DoT-reports), the US (Enron-Andersen) and Germany (ComROAD AG-KPMG).

II.III. THE RESEARCH HYPOTHESIS

As before mentioned, audit failures may downgrade auditor reputation. Subsequently, the impairment of auditor reputation can negatively affect the stock performance of the particular auditor's clients. However, that conclusion is based on international accounting scandals and – as described – results per case are mixed. The existing literature has not focused on Dutch audit failures specifically. This study seeks to fill this gap in literature and assesses the impact of impairment of auditor reputation on the stock performance of the auditor's clients listed on the Dutch Euronext. Since reputation is not directly measurable, we make use of six enforcement initiatives that are imposed by Dutch (financial) authorities as main proxies for the impairment of auditor reputation. The central research question is as follows:

Do the auditor's clients, other than the relevant audited client involved, experience negative abnormal stock returns on the news that their external auditor is penalized by one of the Dutch (financial) authorities?

As the research hypothesis suggests, the main focus is on testing the existence of a potential abnormal market reaction following an event that would have a disparaging impact on the auditor's most fundamental asset, i.e. its reputation. In case this study confirms our hypothesis, i.e. the clients indeed experience negative abnormal stock returns following the event, this would suggest that investors take into account auditor reputation in their investment decisions. In addition, confirmation of the hypothesis would grant additive market power to supervisory authorities, since the impairment of reputation is not limited to the auditor only but is extended to the credibility of the financial statements of the auditor's clients as well. That would lead to the conclusion that the reputation rationale encloses an extra dimension. In that case, auditor reputation does not only affect the auditor itself, but could also have impact on the auditor's clients. As such, appointment of a specific auditor may represent value to the audited company and investors.

III. METHODOLOGY

This section describes our study's methodology and is divided in two parts. Subsection III.I describes the research methodology, whereas subsection III.II elaborates on the quantitative aspects of the research methodology.

III.I. THE DESCRIPTIVE PART

In order to assess the impact of the impairment of auditor reputation on the stock performance of the auditor's clients, we make use of both parametric and nonparametric tests on abnormal stock performance derived on the market model. Parametric tests are generally perceived as the conventional way of conducting event studies, whereas the nonparametric ones are more appropriate in situations that deal with data that depart from normality (Siegel, 1957).[15]

[15] Whether the present study's data are non-normally distributed, is statistically analyzed in section IV.

Prior to conducting the (non)parametric tests, the company's stock prices are transformed into returns and (cumulative) abnormal returns respectively. Both abnormal returns (ARs) and cumulative ARs (CARs) are calculated as the (cumulative) residuals, i.e. the differences between the actual and the expected returns. The expected returns are calculated by the market model that controls for market wide risk (Binder, 1998). The market model is conducted by using the AEX-, AMX- and AScX-index for the companies listed in the respective indices.[16] This results in (C)ARs on individual stock level. By averaging the (C)ARs for both the auditor's clients (TG) and the auditor's non-clients (CG) cross-sectionally, we construct two portfolios in order to measure the overall average impact of the centralized events.[17] By analyzing portfolios instead of individual stocks, we analyze the average impact of the events and remove the effect of firm-specific factors.[18] Moreover, by composing equally weighted portfolios, we control for any cross-correlation problems (Eckbo, 1983; Firth, 1990).

After calculating these portfolios (C)ARs, the values are checked for potential significance using the Student *t*-test. Moreover, we parameterize the (averages of) ARs using an ordinary least squares (OLS) regression framework by regressing the cross-sectionals portfolio's return on an equally weighted market index and dummy variables that are equal to one on particular days during the event window. Hence, the estimated coefficients of the dummy variables represent the portfolio (average of) ARS during the event windows.

Besides these ways of checking the potential significance of the (C)ARs, we perform both parametric and nonparametric *t*-tests in order to analyze whether the average ARs of the TG, experienced during the event windows, are significantly lower than the average ARs of the CG. The differences in means between both groups are tested using several parametric and nonparametric tests, i.e. the (non)paired *t*-test, the Mann Whitney U (MWU) test and the Wilcoxon Sum Rank (WSR) test. The main advantage of the nonparametric tests, relative to their parametric equivalents, is that the first do not require a specific distribution. The nonparametric tests only make use of common sample characteristics, such as the signs and the order of the observations. Hence, nonparametric tests can be applied in almost all types of data. Their main drawback, however, is that nonparametric tests have less power than parametric tests. A potential drawback of the nonparametric tests in the present study, is that the (differences in) time series averages represent the test subject, whereas the parametric tests examine each (C)ARs separately. The advantage thereof, however, is that it focuses on data during a longer time horizon. Hence, it enables us to deal with delay in incorporating news into the stock market.

[16] By using all specific indices instead of using the AEX-index only, we strive to obtain a more accurate market risk beta. The market indices of each events are computed by creating an equally weighted portfolio of all indices.
[17] Although Callen and Morel (2002) decide to match company's stocks from the TG with (comparable) stocks from the CG, we keep all data from both groups and create equally weighted portfolios. The advantage of the present method is that we do not select a limited amount of data. A drawback, however, is that we are bound to work with portfolios (C)ARs instead of comparing individual (C)ARs.
[18] Firm-specific factors, not related to the centralized event(s), have less impact by creating separate portfolios for the TG and CG. By analyzing the impact on the individual stock level, there is a potential risk to generalize conclusions that yield on individual stock level only.

III.II. The Quantitative Part

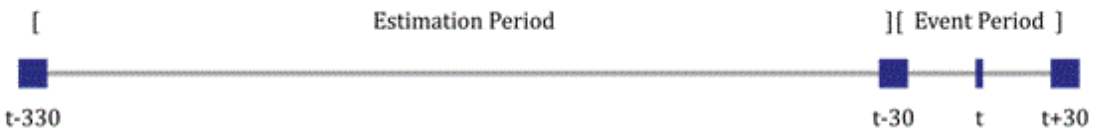
Several (non)parametric analyses are conducted in order to accurately answer the research question. These analyses make use of abnormal returns. The abnormal return (AR) for stock i on a particular day x represents the difference between the actual return of company i on day x, and the estimated normal – or expected – stock return of company i on day x. This is called the market model approach, referring to the returns adjusted for the market wide risk. Hence, stock i's AR on day x, i.e. $AR_{i,x}$, is calculated as residual:

$$AR_{i,x} = R_{i,x} - (\hat{\alpha} + \hat{\beta} \cdot R_{m,x}) \,, \qquad \text{(i)}$$

where $R_{i,x}$ is the actual stock return of company i on day x; $R_{m,x}$ is the return of the market index related to company i on day x; and both $\hat{\alpha}$ and $\hat{\beta}$ are the OLS estimates of the market model parameters.[19] The market model parameters are estimated over the estimation period, which is defined from trading day t-330 to trading day t-31 prior to event day t. These parameters should be estimated prior to the different event windows, since the (potential rumors regarding the) events could affect the estimated $\hat{\beta}$ because of potential event-induced variance or volatility (Binder, 1996; Boehmer, Masumeci & Poulsen, 1991).[20] By choosing an estimation window consisting of 300 observations, potential stationarity problems of the market model parameters are mitigated as well (Blume, 1971; Gonedes, 1973). Moreover, since the lengths of the estimation windows can be considered large, the estimated parameters can expected to be unbiased. That is, the $\hat{\alpha}$ and $\hat{\beta}$ should approach one (1) and zero (0) respectively (Ball & Brown, 1968; Fama, Fisher, Jensen & Roll, 1969; Binder, 1998). Figure II depicts an schematic overview of the way the estimation and event periods are conducted.

Figure II: The Estimation and Event Period

This figure depicts an illustrative overview of the period, covering both the estimation and event period. The market model coefficients are estimated based on the estimation period, covering the time span [t-330, t-31]. Subsequently, these coefficients are used in order to calculate the expected and abnormal returns during the event period, covering the time span [t-30, t+30].



[19] $\hat{\alpha}$ represents the constant, whereas $\hat{\beta}$ is the coefficient for market risk. Hence, the market model coefficients mitigate the risk that economy wide factors influence stock returns (Binder, 1998).

[20] Event-induced variance (volatility) refers to the potential increase in variance (volatility) when the announcement of news conveys relevant information (Boehmer, Musumeci & Poulsen, 1991).

Based on the values of these ARs, the value of the CAR during a particular event window [x, y] is easily calculated by summing up the individual ARs from day x to day y:

$$CAR_{i,[x,y]} = \sum_{n=x}^{y} AR_{i,n}, \quad \text{(ii)}$$

Since there is no convention on the number of days that should be included, we analyze the daily ARs and CARs during event windows of different lengths. Since news is incorporated into equity prices gradually, we calculate daily ARs and CARs for the event windows.

III.II.I. The Analysis on (C)ARs

According to the literature, parametric tests are the conventional way of testing ARs. In order to measure the impact of the centralized events on the stock performance, the parametric tests are performed in two ways. First of all, the potential significance of both the TG's and the CG's (C)ARs, experienced during both a 6-day period [t, t+5] and 31-day period [t, t+30] are analyzed. This subsection specifies on the methodology of testing the (significance of the) individual (C)ARs.

First, the ARs are checked for potential significance. The significance of the portfolios ARs is calculated by dividing the value of the AR to its estimated standard deviation:

$$t - \text{test statistic} = \frac{AR_x}{\hat{s}(AR)}, \quad \text{(iii)}$$

where AR_x denotes the portfolio AR on particular day x, as calculated in the way described in equation (i); and $\hat{s}(AR)$ represents an estimate of the standard deviation, calculated over the 300-day estimation period, covering the period [t-330, t-31]. The estimate of the standard deviation is calculated as described in equation (iv).

$$\hat{s}(AR) = \sqrt{\frac{1}{T_0 - 1} \cdot \sum_{x=t-330}^{t-31} [AR_x - \overline{AR}]^2}, \quad \text{(iv)}$$

where T_0 represents the number of trading days during the estimation period (i.e. 300); and $\overline{AR}$ denotes the time series average of all ARs during the estimation period. The calculated test statistic follows a Student's t-distribution with (T_0-1) degrees of freedom, which is assumed distributed standard normal under the null hypothesis of no abnormal performance. By calculating the standard deviation over the estimation period, the parameter is not affected by the event (Brown & Warner, 1980).

The significance of the CAR during a particular period [x, y] is calculated in a similar manner. The significance is calculated by dividing the CAR by its estimated standard deviation, which is presented mathematically in equation (v).

$$t - \text{test statistic} = \frac{CAR_{[x,y]}}{\hat{s}(CAR_{[x,y]})}, \quad \text{(v)}$$

where the $CAR_{[x,y]}$ is calculated as described in equation (ii); and $\hat{s}(CAR_{[x,y]})$ represents its estimated standard deviation and is calculated as the square root of the sum of the variance estimates during the particular event window [x, y]:

$$\hat{s}(CAR_{[x;\,y]}) = \sqrt{\sum_{n=x}^{y} \hat{s}(\overline{AR_n})^2}\,, \qquad \text{(vi)}$$

where $\hat{s}(\overline{AR_n})^2$ is defined as the square of the standard deviation estimate of equation (iv).

In order to check the robustness of (the significance of) the aforementioned (C)ARs, an OLS regression framework measures the effects of the events as well. The (average) ARs during both event windows are modeled as coefficients in an OLS regression model. The dependent variable covers the (average) realized returns during both the estimation and the event period for each portfolio. This dependent variable is regressed on both an equally weighted market beta and some dummy variables for the days following the event. Following Binder (1998) and Callen and Morel (2002), we use an equally weighted portfolio for the actual, realized returns – hence, not the ARs – as the dependent variable, and a single equation is regressed as described in equation (vii).

$$R_{i,x} = \propto_i + \beta_i \cdot R_{i_m,x} + \sum_{t=T1}^{T2} \cdot c_{i,x} \cdot D_{i,x} + \varepsilon_{i,x}\,, \qquad \text{(vii)}$$

where $R_{i,x}$ is the time series of the equally weighted return of portfolio i, consisting of the average return of either the TG or the CG, on day x; [21] $R_{i_m,t}$ is the market return of the equally weighted portfolio of the TG or CG on day x; and $D_{i,x}$ is a dummy variable that equals one (1) at a particular event day x. The estimated coefficients $\propto_i$, β_i and $c_{i,x}$ are respectively the intercept of the market model, the slope of the market model, and the estimate of the daily overall average AR on a particular event day x. Besides the dummy variables that parameterize the different ARs, the average ARs during an event windows is parameterized by adding a dummy variable that equals one (1) for all days covering the particular event window $[x, y]$.[22]

The underlying rationale for this OLS regression model is that it parameterizes the overall (average) market reaction, which cannot be estimated with the market model parameters only, at different event days (windows) during the event period. The potential significance of these (average) ARs are tested using the standard t-test. The underlying null hypotheses suggest that the estimated (average) ARs are not significantly different from zero (0). This suggests that in case the null hypothesis is not rejected, the particular portfolio return is not significantly different from its market model expectations.

[21] An alternative to measure the impact of the event is to specify the OLS regression equation for each firm separately.
[22] In short, the dummy variables for event windows [t, t+5] and [t, t+30] equals one during the 6- and 31-day event window respectively. Hence, the coefficients represent the average ARs during the particular event window.

Testing whether the null hypothesis can be rejected or not, the significance of the estimated coefficients on a particular day during the event period, is measured with the t-test statistic as presented in equation (viii).

$$t - test\ statistic = \frac{\hat{x}}{\sqrt{\frac{SD(\hat{x})^2}{N}}}\ , \tag{viii}$$

III.II.II. The Analysis on Differences

In addition, both parametric and nonparametric tests are performed in order to give a more accurate conclusion on the question whether the portfolios averages differ significantly from each other.

The parametric (un)paired t-tests are based on the common null hypothesis that both portfolios are equal (in averages):

$$H_0: \overline{AR}_{TG} = \overline{AR}_{CG}\ , \tag{ix}$$

where $\overline{AR}_{TG}$ ($\overline{AR}_{CG}$) represents the average of the portfolios' ARs during the several event windows for the TG (CG). This null hypothesis is tested against the one sided alternative hypothesis (H_1) that the time series average of the TG is lower than its counterpart: $\overline{AR}_{TG} < \overline{AR}_{CG}$. This hypothesis is checked by using the paired and unpaired variants of the t-tests.[23]

First, the paired t-test statistics are calculated by computing the differences between both portfolios ARs for each day during the event windows.[24] After calculating these differences, i.e. D_x, the test statistics of the paired t-test are computed as described in equation (x).

$$paired\ t - test\ statistic = \frac{\overline{D}}{\sqrt{\frac{\hat{s}(D)^2}{N}}}\ , \tag{x}$$

where $\overline{D}$ is the time series average of the differences between both portfolios during the particular event window; $\hat{s}(D)$ ($\hat{s}(D)^2$) is the standard deviation (variance) of the differences between both portfolios during the event window; and N represents the number of days during the event window .[25]

The unpaired t-test statistics are calculated differently. Since the unpaired t-test is used for testing two independent samples, it makes use of the time series mean and the standard deviation of each portfolios instead of (the parameters based on) the differences between the portfolios. The unpaired t-test statistic is calculated as described in equation (xi).

$$unpaired\ t - test\ statistic = \frac{\overline{AR_{CG}} - \overline{AR_{TG}}}{\sqrt{\frac{[\hat{s}(AR_{CG})^2 \cdot (N_{CG}-1)] + [\hat{s}(AR_{TG})^2 \cdot (N_{TG}-1)]}{N_{CG} \cdot N_{TG}}}}\ , \tag{xi}$$

[23] Although both t-tests assume that the analyzed data is from a normal distribution, the tests differ from each other. The main difference is that the paired t-test requires two samples that are equal, whereas this is not required in case of the unpaired variant. Generally, the paired t-test tests for differences between the subjects at two points of time, whereas the unpaired t-test can perfectly deal with independent – hence, different – samples. In the present study both tests are performed.

[24] The differences between both portfolios, denoted by D_x, are calculated as follows: $D_x = AR_{CG,x} - AR_{TG,x}$

[25] Hence, N is equal to 6 and 31 in case of the 6- and 31-day event window respectively.

where $\overline{AR_i}$ is the time series mean of the portfolios ARs of portfolio i during a particular event window; $\hat{s}(AR_i)$ ($\hat{s}(AR_i)^2$) denotes the estimated standard deviation (variance) of portfolio i during a particular event window; and N_i represents the number of days of portfolio i.

Since estimates of the standard deviations are used, the paired t-test statistic follows a Student's t-distribution with (N_D-1) degrees of freedom. The unpaired t-test statistic, however, follows a different distribution with $(N_{CG} + N_{TG} - 2)$ degrees of freedom. In case the t-test statistics exceed the critical value of the Student's t-distribution, the null hypothesis is rejected.

Besides the parametric t-tests, both their nonparametric equivalents, the MWU test and the WSR test, are performed in order to check the robustness of the results of the t-tests. Both the MWU and the WSR tests detect whether two samples stem from the same population. The differences between the tests are, however, that the WSR test examines whether there are differences between paired samples, whereas the MWU test is based on the differences between two unpaired – independent – samples.[26] Since both tests examine whether the samples stem from an the same population, the tests yield the following common null hypothesis:

$$H_0: \overline{AR}_{TG} = \overline{AR}_{CG}\,, \tag{xii}$$

where $\overline{AR}_{TG}$ ($\overline{AR}_{CG}$) represents the time series' average of the TG (CG) during the different event windows. This null hypothesis is tested against the directive variant (H_1) yielding that the time series average of the TG is lower than its counterpart: $\overline{AR}_{TG} < \overline{AR}_{CG}$.

For the paired WSR tests all absolute differences in ARs during the event windows should be pooled and ranked first. Then, after adding minus signs to negative differences, the sum of ranks are calculated for both the positive and negative differences separately. The conclusion whether the TG performs worse relative to the CG, is drawn based on the comparison of these sum of ranks. That is, if the event has no effect, the sum of ranks should be similar, whereas the sum of ranks differ in case the event would have any effect. In short, the sum of both ranks can be calculated as mathematically presented in equation (xiii).

$$\begin{aligned} T_+ &= \sum r\,, \\ T_- &= \sum -r\,, \end{aligned} \tag{xiii}$$

where r represents the ranks of the differences between the paired observations and T_+ (T_-) depicts the sum of the positive (negative) signed ranks. Subsequently, the WSR test statistic, i.e. $WSR = min(T_+, T_-)$, is compared to the critical value of the WSR test in order to conclude whether to reject the underlying null hypothesis, implying there is no (significant) difference between the two portfolios. In case the WSR test statistic is below its critical value, the null hypothesis should be rejected.

[26] Hence, the difference between the WSR and the MWU tests is directly comparable to the difference between the paired and unpaired t-tests.

In order to perform the MWU test, all differences between the ARs for both portfolios are pooled during the event periods. Subsequently, the ARs for the TG and CG together are ranked in increasing order, with the lowest AR having rank 1 and the highest AR having rank either 12 or 62, depending on the length of the chosen event window. The MWU test statistics could then be calculated for both the TG and the CG in the way described in equation (xiv).

$$MWU_i = N_{TG} \cdot N_{CG} + \frac{1}{2} \cdot N_i(N_i + 1) - R_i \,, \tag{xiv}$$

where N_{TG} (N_{CG}) and R_i represents the number of ARs of the TG (CG) and the sum of ranks of portfolio i, consisting of either one of the two portfolios respectively. Subsequently, the MWU test statistic, i.e. $MWU = min(MWU_{TG}, MWU_{CG})$, is compared to the critical value of the MWU test in order to conclude whether to reject the null hypothesis. In case the MWU test statistics is below its critical value, the null hypothesis, implying there is no (significant) difference between the two portfolios, should be rejected.

IV. DATA

This section introduces the centralized data in both a descriptive and statistical way. Subsection IV.I describes the centralized events as well as the used data. Subsequently, subsection IV.II presents a statistical analysis thereof.

IV.I. THE DESCRIPTION OF THE DATA

The present study centralizes six events, mainly comprising of news publications on enforcement sanctions imposed by the Dutch Authority for the Financial Markets (AFM), i.e. the financial supervisor on the Dutch audit market, and the Dutch public prosecutor. These events serve as proxies of the impairment of the reputation of the particular auditor involved. These events receive a certain amount of media attention, in which the name of the relevant auditor as well as the reason why the specific auditor is sanctioned are both mentioned.

We specify the event dates as the day on which the news reaches the market that the supervisory authority imposes the enforcement sanction. We assume that the day the market has processed the news is equal to the day on which these sanctions are published on the financial authority's website.[27] Furthermore, it is essential that (i) the enforcement initiative is published and – thus – publicly available for investors on the stock market, (ii) the involved auditor is named in this news publication and (iii) the involved auditor is the (external) auditor of at least one of the companies listed on the centralized indices of the Euronext Amsterdam. The centralized events are briefly described in table I.

[27] I refer to the websites of the authorities involved.

Table I: The Events Described

This table briefly introduces the six main events. Sequentially, this table includes the date of the event, its cited name and a short description of the particular event. The events are presented in chronological order.

Event date	Cited name	Description of the event
Jan 26, 2012	EY-I	The AFM's decision to impose a EUR 54,450.- fine is official and published.
Feb 23, 2012	DELOITTE-I	The AFM's decision to impose a EUR 54,450.- fine is official and published.
Oct 02, 2012	EY-II	The AFM's decision to impose a EUR 217,810.- fine is official and published.
Jun 06, 2013	KPMG-I	The AFM's decision to impose a EUR 881,250.- fine is official and published.
Aug 19, 2013	KPMG-II	The Accountantskamer's decision to issue a reprimand is official and published.
Dec 30, 2013	KMPG-III	The transaction with the Dutch prosecutor of EUR 7,000,000.- for the Ballast Nedam case.

Although each of the three indices of the Euronext Amsterdam consist of 25 companies, we acknowledge the presence of inactive companies, i.e. the companies with a low liquidity and a constant return of 0.00%, and the companies that are not listed during (a part of) the estimation period. By correcting for these companies, we mitigate for the (potential) risk that these companies could influence outcomes.[28] Appendix A depicts an overview of the composition of the different portfolios.[29]

Subsequently, the companies' stock prices are gathered from Datastream. These stock prices are converted into stock returns per individual company first.[30] The data consists of 361 observations per company, covering both the estimation and event period of [t-330, t+30] around event date t. Hence, the total dataset consists of an initial (75 · 361 =) 27,075 firm-trading day observations per individual event. All events combined, the original dataset comprises (6 · 27,075 =) 162,450 firm-trading day observations. We omitted 25 inactive companies as per the criteria above, which results in a final dataset including 153,425 firm-trading day observations divided for the six centralized events. Based on these firm-trading day observations, (the (C)ARs of) the portfolios are constructed as according the methodology described in section III. The cumulative returns of each portfolio, covering the total period [t-330, t+30] and the event period [t-30, t+30] are presented in appendix B. These graphical representations indicate that the TG in case of EY-I, DELOITTE-I and KPMG-I seems to perform substantially worse during the first few days following the event date t. However, it is no so straightforward to draw robust inferences on these graphs, since the (cumulative) daily returns are neither controlled for market wide risk nor tested for potential statistical significance. Therefore, the portfolios' (C)ARs during the event windows as well as the differences between the TG and CG during these periods are centralized and tested for significance using

[28] These companies, comprising of (i) companies with non-traded stocks and (ii) partially listed companies, could bias final results, since these companies yield zero (0) (abnormal) returns for at least part of the estimation period. As such, these companies may give an incorrect – or biased – upward (downward) pressure on negative (positive) portfolio averages, which could to a somewhat biased conclusion. Therefore, these observations are manually omitted from the database.

[29] Data on the company's external auditor is gathered from Company.info and manually checked by the companies' annual reports of the respective year. We centralize the specific date the sanction is published, e.g. the TG (CG) of EY-I consists of the publicly listed companies that were audited (not audited) by EY during the year 2012. The companies relevant indices are gathered from iex.nl and aex.nl.

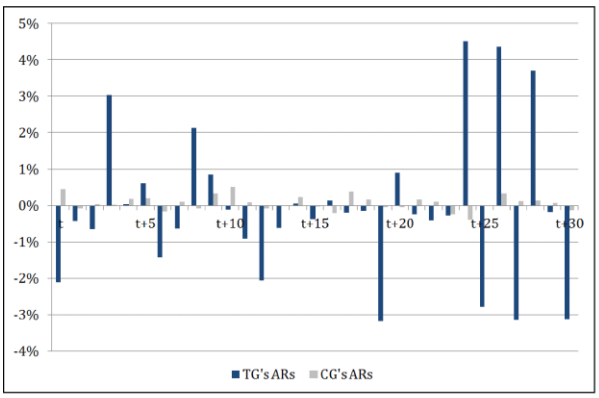
[30] Dividing the (for dividends adjusted) closing stock prices of a particular company at time $t-1$ by its stock prices at time t and subtracting it with 1, results in the company's stock return at time t.

the aforementioned tests. The portfolios (C)ARs, centralized as research subject of the present study, are presented in figure III.

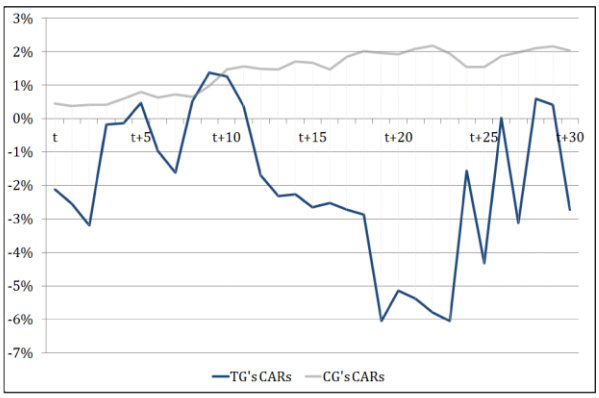
Figure III: (Cumulative) Abnormal Returns – Part I

The left figures depict graphical representations of the abnormal returns during the event period for each event separately. The right figures depict graphical representations of the cumulative abnormal returns during the event for each event separately.

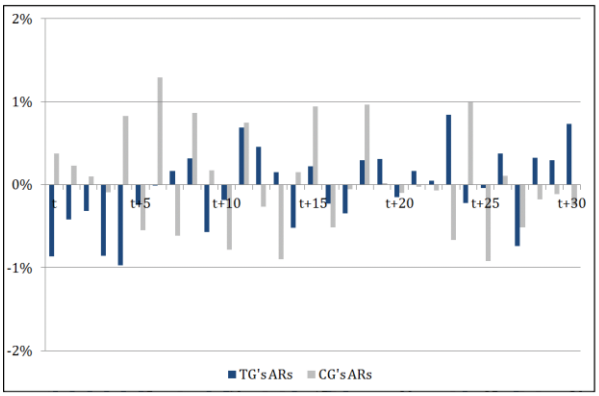
EY-I



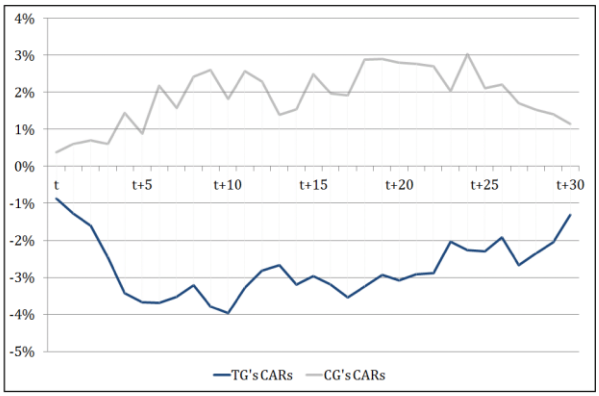
EY-I



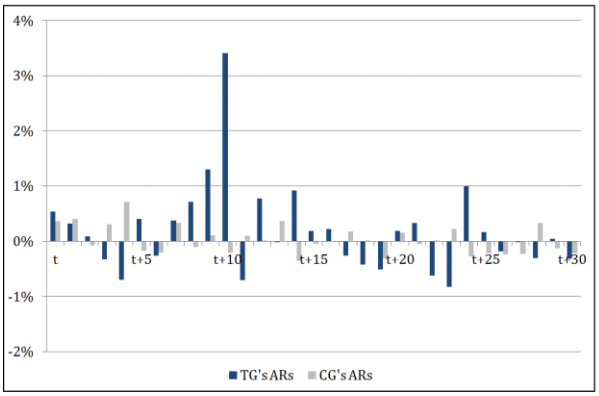
Deloitte-I



Deloitte-I



EY-II



EY-II

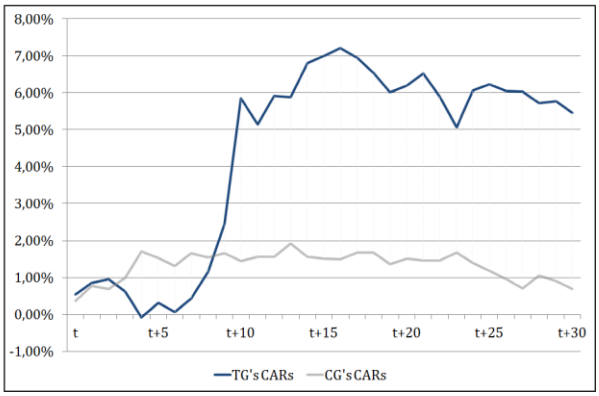
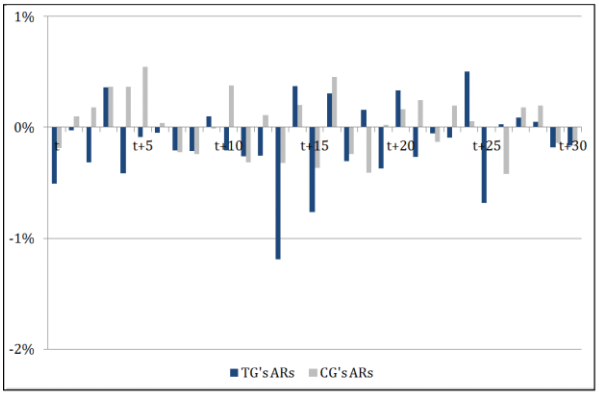
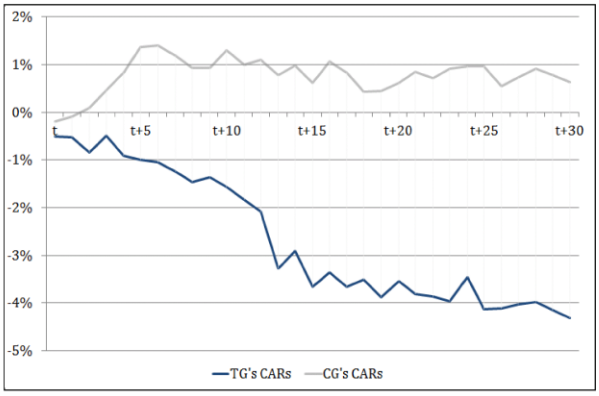


Figure III: (Cumulative) Abnormal Returns – Part II

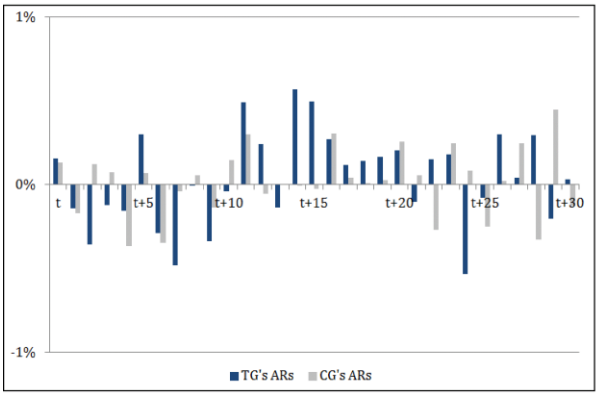
KPMG-I



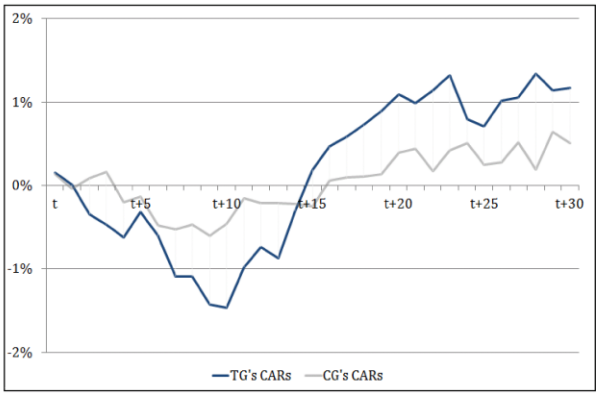
KPMG-I



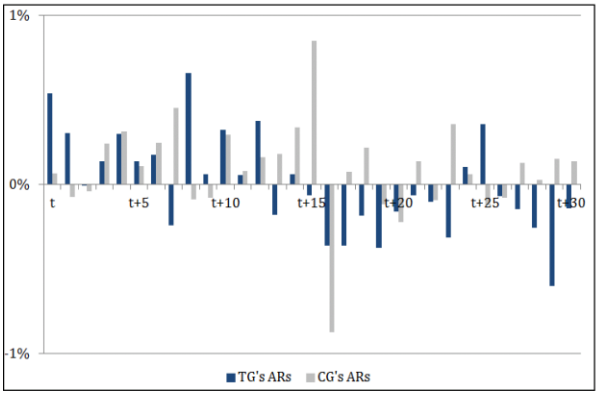
KPMG-II



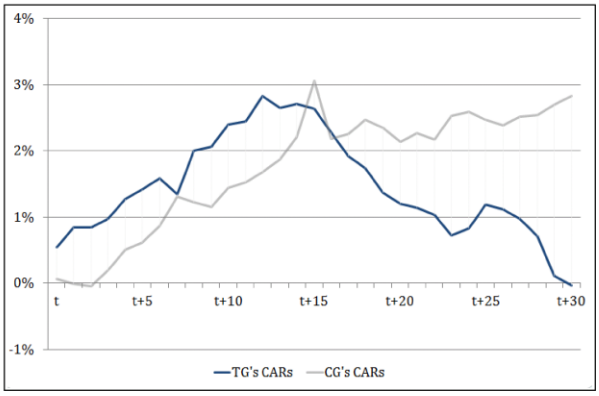
KPMG-II



KPMG-III



KPMG-III



IV.II. THE ANALYSIS OF THE DATA

As indicated by earlier studies, (abnormal) stock returns on daily basis are often not distributed normally (e.g. Fama, 1976; Brown & Warner, 1985). These studies suggest that daily (abnormal) returns are relatively fat-tailed, compared to the normal distribution. As presented in table II, also the degree of skewness of this study's (abnormal) returns deviates from the normal distribution. Furthermore, table II depicts the other descriptive statistics of the centralized data, including both the aforementioned degree of skewness and (excess) kurtosis.

Hence, the values of table II indicates that the samples are non-normally distributed, unless table II indicates that the portfolios means and medians of the (abnormal) returns are located around 0.00%. The majority of positive skewness values indicates that the distributions are right skewed, suggesting that the sample are asymmetrically distributed and the majority of the values is located left from the sample's average. Furthermore, in line with earlier literature, the values of the (excess) kurtosis indicates that the data have relative fat tails. Moreover, the highly significant Jarque-Bera test statistics of the daily (abnormal) returns, which is implicitly based on the measures of (excess) kurtosis and skewness, indicate that the underlying null hypothesis of normality should be rejected.[31] Hence, based on the aforementioned aspects and as already suggested by previous literature, the portfolios' (abnormal) returns are not distributed normally. This implies that parametric tests suffer from limitations, since at least one of the parametric statistics' assumptions is not satisfied.[32] Since the majority of the portfolios is positively skewed, parametric tests reject too seldom (too often) when testing for negative (positive) abnormal performance. Nonparametric tests have a wider application, since these tests do not require all criteria of parametric tests (Serra, 2002).

Besides the check on the characteristics of the dependent variables, the time series variables are checked for stationarity as well. If the variables in the multivariate regression model are not stationary, i.e. have a unit root, the time series' mean and variance are not constant over time. In that case the *t*-test statistics do not follow a Student's *t*-distribution, which yield consequences for interpreting the parametric test results. Table II depicts the results of the different kind of unit root tests.[33] When results of the unit root tests are significant, the underlying null hypotheses that the variable is non-stationary are rejected in all cases. Based on these tests results, it can be concluded that all individual time series are stationary and do not follow any trend over time. In short, there are no potential issues in interpreting the *t*-test statistics in terms of non-stationarity.

[31] Next to the performed Jarque-Bera test, also the unreported Shapiro-Wilk test statistics indicated that the underlying null hypothesis that the time series are normally distributed, should be rejected at 1% significance levels.
[32] Parametric statistics require that (i) the population is normally distributed, (ii) the variances are equal among the population, (iii) the observations are independent and (iv) the variables are measured at least at interval scale.
[33] The (equally weighted) portfolios of each market index are checked for stationarity as well. Although these test results are not reported, all test statistics suggest that the null hypotheses of non-stationarity are rejected.

Table II: Descriptive Statistics and Unit Root Tests

This table depicts relevant descriptive statistics of the time series' daily (abnormal) returns ((A)R) for both the treatment group (TG) and control group (CG). All time series consist of 361 observations, covering both the full estimation and event period. The mean, median, standard deviation, skewness, Kurtosis and Jarque-Bera test statistics are presented for each individual portfolio. Moreover, the results of the Augmented Dickey Fuller (ADF) tests and the Philips-Perron (PP) tests are depicted. In case of the latter tests, potential significance can be interpreted as described in footnote 1.

Event name	Return Type	Descriptive statistics												Unit Root Tests			
		Obs.	*Mean*		*Median*		*Std. deviation*		*Skewness*		*Kurtosis*		*Jarque-Bera*		*ADF*		*PP*	
		TG / CG	*TG*	*CG*	*TG*	*CG*	*TG*	*CG*	*TG*	*CG*	*TG*	*CG*	*TG*	*CG*	*TG*	*CG*	*TG*	*CG*
EY-I	*R*	*361*	*-0.0001*	*-0.0002*	*0.0000*	*0.0001*	*0.0188*	*0.0128*	*-0.1525*	*-0.3984*	*5.1141*	*5.1290*	*68.62* ***	*77.73* ***	*-19.35* ***	*-16.29* ***	*-19.37* ***	*-16.26* ***
	AR	*361*	*0.0000*	*0.0001*	*-0.0011*	*0.0001*	*0.0108*	*0.0025*	*0.5712*	*0.0119*	*5.5858*	*5.4805*	*120.20* ***	*92.56* ***	*-17.27* ***	*-20.44* ***	*-29.70* ***	*-20.44* ***
DELOITTE-I	*R*	*361*	*-0.0001*	*-0.0002*	*0.0006*	*0.0000*	*0.0132*	*-0.3918*	*-0.3477*	*0.0138*	*4.2454*	*5,2215*	*30.60* ***	*83.47* ***	*-16.27* ***	*-11.17* ***	*-16.30* ***	*-16.54* ***
	AR	*361*	*0.0000*	*0.0001*	*0.0001*	*-0.0001*	*0.0049*	*0.0036*	*0.4537*	*0.2717*	*6.0126*	*4.0597*	*148.90* ***	*21.33* ***	*-17.04* ***	*-23.76* ***	*-17.05* ***	*-24.25* ***
EY-II	*R*	*361*	*-0.0006*	*-0.0003*	*-0.0003*	*0.0002*	*0.0194*	*0.0138*	*-0.0445*	*-0.3374*	*4.0284*	*4.4268*	*16.03* ***	*37.47* ***	*-16.92* ***	*-16.57* ***	*-16.81* ***	*-16.42* ***
	AR	*361*	*0.0002*	*0.0000*	*-0.0001*	*0.0001*	*0.0070*	*0.0028*	*0.5814*	*-0.1239*	*6.0967*	*7.7873*	*164.59* ***	*345.65* ***	*-17.76* ***	*-18.43* ***	*-17.72* ***	*-18.48* ***
KPMG-I	*R*	*361*	*0.0002*	*0.0003*	*0.0003*	*0.0002*	*0.0091*	*0.0099*	*-0.6032*	*-0.2135*	*4.8337*	*4.6171*	*72.47* ***	*42.08* ***	*-17.71* ***	*-17.14* ***	*-17.67* ***	*-17.06* ***
	AR	*361*	*-0.0000*	*0.0002*	*-0.0000*	*0.0001*	*0.0040*	*0.0043*	*0.1948*	*1.7912*	*6.7490*	*20.3797*	*213.69* ***	*4,736.42* ***	*-17.45* ***	*-17.55* ***	*-17.45* ***	*-17.54* ***
KPMG-II	*R*	*361*	*0.0005*	*0.0007*	*0.0008*	*0.0009*	*0.0081*	*0.0087*	*-0.5232*	*-0.2744*	*5.1785*	*4.4375*	*87.85* ***	*35.61* ***	*-17.15* ***	*-11.85* ***	*-17.15* ***	*-18.98* ***
	AR	*361*	*0.0001*	*0.0001*	*0.0001*	*-0.0000*	*0.0039*	*0.0034*	*0.0861*	*-0.5512*	*6.8279*	*6.5645*	*220.85* ***	*209.39* ***	*-17.42* ***	*-16.18* ***	*-17.41* ***	*-15.98* ***
KMPG-III	*R*	*361*	*0.0005*	*0.0010*	*0.0004*	*0.0015*	*0.0070*	*0.0072*	*-0.7892*	*-0.5937*	*6.2977*	*4.1806*	*201.05* ***	*42.17* ***	*-17.07* ***	*-17.29* ***	*-17.15* ***	*-17.26* ***
	AR	*361*	*0.0009*	*0.0010*	*0.0003*	*0.0008*	*0.0117*	*0.0127*	*0.0181*	*0.1509*	*4.3897*	*4.1545*	*29.07* ***	*21.42* ***	*-16.45* ***	*-16.55* ***	*-16.41* ***	*-16.46* ***

*[1] In case of the ADF- and PP-tests, the significance at a significance level of 10%, 5% and 1% is denoted with *, **, ***, respectively.*

V. RESULTS

This section discusses the empirical results. Firstly, subsection V.I focuses on the individual (C)ARs and centralizes the question whether either the TG or the CG experiences a significant (negative) market reaction following the events. Subsequently, subsection V.II analyzes the differences between both groups by discussing the test results of both parametric and nonparametric tests and assesses the hypothesis that the TG yields significantly lower abnormal returns than the CG following the events. Finally, subsection V.III summarizes the results.

V.I. THE (C)ARS

This subsection mainly focuses on the portfolios (C)ARs following the different events. Both table III and IV presents the results of the parametric tests performed. Table III depicts an overview of the experienced (C)ARs, covering the broadest event window [t, t+30]. Table IV presents the results of the estimated OLS regression models. Table III (table IV) should be read as follows: the percentages (decimals) denote the sign and magnitude of the (C)ARs, whereas the added asterisks represent potential significance.[34] Based on these test results, the (C)ARs of the TGs and the CGs are discussed.[35] As depicted in table IV, the models have large values of (adjusted) R^2.[36] Moreover, all models are statistically significant at a 1% significance level, according to the overall F-test statistics.[37] This subsection's remainder focuses on the ARs on the different days. The market model coefficients are not discussed in detail.[38]

[34] E.g. in case of EY-I, the TG's AR on trading day t is significantly negative at a 1% significance level: -2.12%. Its counterpart, i.e. the CG's AR on that day, is significantly positive at a 10% significance level: 0.44%.
[35] This section discusses the results as depicted in table III. However, in case the results of the OLS regression models in table IV differ significantly from the results of table III, this will be discussed as well.
[36] The model's R^2 represents the proportion of variance in the dependent variable that can be explained by the independent variable(s). The adjusted R^2 corrects for addition of (an) extraneous independent variable(s).
[37] Since all test statistics are highly significant, the results of these overall F-tests are not reported.
[38] In all cases, the estimated coefficients of $\propto_P$, the intercept of the estimated regression model, is not significant at any conventional significance level. Thus, it can be concluded that there is no evidence to reject the null hypothesis that $\propto_P$ is equal to zero (0). In contrast, the estimated coefficients of β_P are all statistically significant at a 1% significance level.

Table III: The Individual (C)ARs analysis

This table depicts the portfolio (C)ARs during the several event windows for each event seperately. Potential significance in case of the treatment group (TG) can be interpreted as described in footnote 1, whereas potential significance in case of the control group (CG) can be interpreted as described in footnote 2.

Panel A: The Individual (C)AR Analysis

EY-I

	TG				CG			
Day	AR	S	CAR	S	AR	S	CAR	S
t	-2,12%	***	-2,12%	**	0,44%	*	0,44%	*
t+1	-0,42%		-2,54%	**	-0,07%		0,37%	
t+2	-0,66%		-3,20%	**	0,04%		0,41%	
t+3	3,02%	***	-0,17%		0,00%		0,42%	
t+4	0,03%		-0,14%		0,18%		0,60%	
t+5	0,60%		0,46%		0,19%		0,79%	
[t+6, t+10]	0,81%		1,27%		0,69%		1,48%	*
[t+11, t+20]	-6,42%		-5,15%		0,44%		1,92%	*
[t+21, t+30]	2,43%		-2,72%		0,11%		2,03%	

DELOITTE-I

	TG				CG			
Day	AR	S	CAR	S	AR	S	CAR	S
t	-0,86%	**	-0,86%	**	0,37%		0,37%	
t+1	-0,42%		-1,29%	**	0,23%		0,60%	
t+2	-0,32%		-1,60%	**	0,10%		0,70%	
t+3	-0,85%	**	-2,46%	***	-0,09%		0,61%	
t+4	-0,97%	**	-3,43%	***	0,83%	**	1,44%	**
t+5	-0,25%		-3,67%	***	-0,55%	*	0,89%	
[t+6, t+10]	-0,30%		-3,97%	***	0,93%		1,82%	*
[t+11, t+20]	0,88%		-3,09%	*	0,98%		2,80%	*
[t+21, t+30]	1,77%		-1,31%		-1,65%		1,15%	

EY-II

	TG				CG			
Day	AR	S	CAR	S	AR	S	CAR	S
t	0,54%		0,54%		0,36%		0,36%	
t+1	0,32%		0,86%		0,40%		0,77%	*
t+2	0,09%		0,95%		-0,08%		0,69%	
t+3	-0,32%		0,62%		0,30%		0,99%	*
t+4	-0,70%		-0,08%		0,71%	**	1,70%	***
t+5	0,40%		0,32%		-0,18%		1,53%	**
[t+6, t+10]	5,53%		5,85%		-0,08%		1,45%	
[t+11, t+20]	0,35%		6,20%		0,06%		1,51%	
[t+21, t+30]	-0,74%		5,46%		-0,82%		0,69%	

KPMG-I

	TG				CG			
Day	AR	S	CAR	S	AR	S	CAR	S
t	-0,51%	*	-0,51%	*	-0,18%		-0,18%	
t+1	-0,03%		-0,53%		0,10%		-0,09%	
t+2	-0,32%		-0,85%		0,18%		0,09%	
t+3	0,36%		-0,49%		0,37%		0,46%	
t+4	-0,42%		-0,90%		0,37%		0,82%	
t+5	-0,09%		-0,99%		0,54%		1,37%	
[t+6, t+10]	-0,58%		-1,57%		-0,06%		1,31%	
[t+11, t+20]	-1,98%		-3,55%	**	-0,70%		0,61%	
[t+21, t+30]	-0,77%		-4,31%	**	0,03%		0,64%	

KPMG-II

	TG				CG			
Day	AR	S	CAR	S	AR	S	CAR	S
t	0,16%		0,16%		0,13%		0,13%	
t+1	-0,14%		0,01%		-0,17%		-0,04%	
t+2	-0,36%		-0,34%		0,12%		0,09%	
t+3	-0,12%		-0,47%		0,08%		0,16%	
t+4	-0,15%		-0,62%		-0,37%		-0,20%	
t+5	0,30%		-0,32%		0,07%		-0,14%	
[t+6, t+10]	-1,15%		-1,47%		-0,32%		-0,46%	
[t+11, t+20]	2,56%		1,09%		0,85%		0,39%	
[t+21, t+30]	0,08%		1,17%		0,12%		0,51%	

KPMG-III

	TG				CG			
Day	AR	S	CAR	S	AR	S	CAR	S
t	0,54%		0,54%		0,06%		0,06%	
t+1	0,30%		0,84%		-0,07%		-0,01%	
t+2	0,00%		0,84%		-0,04%		-0,05%	
t+3	0,14%		0,98%		0,24%		0,19%	
t+4	0,30%		1,28%		0,31%		0,50%	
t+5	0,14%		1,41%		0,11%		0,61%	
[t+6, t+10]	0,98%		2,39%		0,83%		1,44%	
[t+11, t+20]	-1,19%		1,20%		0,69%		2,13%	
[t+21, t+30]	-1,23%		-0,03%		0,70%		2,83%	

*[1] The significance at a one sided significance level of 10%, 5% and 1% is denoted with *, **, ***, respectively.*
*[2] The significance at a two sided significance level of 10%, 5% and 1% is denoted with *, **, ***, respectively.*

Table IV: OLS Regression Models - Part I

This table depicts the output of the OLS regression models. Panel A (Panel B) reports the estimates of the OLS regression coefficients for the treatment group (control group). The estimated coefficients of the alpha, beta and the relevant dummy variables are reported as well as its adjusted R^2. Potential significance can be interpreted as described in footnote 1.

Panel A: The Treatment Group: The Auditor's Clients

Event date	Event name		Estimated OLS coefficients										Adj. R^2
			α	β	t	$t+1$	$t+2$	$t+3$	$t+4$	$t+5$	$[t, t+5]$	$[t, t+30]$	
26-01-2012	EY-I	TG	0,0001 (0.20)	1.2547*** (26.50)	-0.0210* (-1.94)	-0.0062 (-0.57)	-0.0064 (-0.59)	0.0302*** (2.78)	-0.0006 (-0.06)	0,0055 (0.51)			0,6675
		TG	0.0001 (0.20)	1.2563*** (26.47)							0.0002 (0.05)		0,6605
		TG	0.0002 (0.36)	1.2574*** (26.52)								-0.0012 (-0.57)	0,6608
23-02-2012	Deloitte-I	TG	0.0002 (0.75)	1.0174*** (48.48)	-0.0088* (-1.82)	-0,0044 (-0.92)	-0,0032* (-0.67)	-0.0087* (-1.81)	-0,0098** (-2.04)	-0.0026 (-0.53)			0.8679
		TG	0.0002 (0.76)	1.0183*** (48.78)							-0.0063*** (-3.18)		0.8689
		TG	0.0001 (0.49)	1.0181*** (48.11)								-0.0005 (-0.53)	0.8654
02-10-2012	EY-II	TG	-0.0003 (-0.72)	1.3655*** (47.35)	0.0052 (0.71)	0.0032 (0.44)	-0.0001 (-0.02)	-0.0034 (-0.47)	-0.0081 (-1.12)	0.0034 (0.48)			0.8617
		TG	-0.0003 (-0,73)	1.3651*** (47.57)							0.0000 (0.01)		0.8627
		TG	-0.0004 (-1.06)	1.3648*** (47.66)								0.0017 (1.24)	0.8632
03-06-2013	KPMG-I	TG	-0.0001 (-0.42)	1.009*** (38.85)	-0.0049 (-1.24)	-0.0005 (-0.12)	-0.0027 (-0.68)	0.0038 (0.96)	-0.0048 (-1.20)	-0.0008 (-0.20)			0.8090
		TG	-0.0001 (-0.42)	1.0075*** (39.08)							-0.0017 (-1.01)		0.8099
		TG	0.0001 (0.07)	1.0090*** (39.35)								-0,0015** (-2.05)	0.8115
19-08-2013	KMPG-II	TG	-0.0000 (-0.19)	0.9901*** (34.59)	0.0008 (0.19)	-0.0012 (-0.31)	-0.0036 (-0.91)	-0.0015 (-0.38)	-0.0022 (-0.56)	0.0034 (0.86)			0.7691
		TG	-0.0000 (-0.19)	0.9894*** (34.83)							-0.0007 (-0.45)		0.7711
		TG	-0.0001 (-0.32)	0.9896*** (34.84)								0.0002 (0.30)	0.7710
30-12-2013	KPMG-III	TG	-0.0002 (-1.03)	0.9618*** (29.79)	0.0053 (1.41)	0.0033 (0.87)	0.0002 (0.06)	0.0023 (0.63)	0.0024 (0.63)	0.0014 (0.37)			0.7118
		TG	-0.0002 (-1.04)	0.9623*** (30.05)							0.0025 (1.62)		0.7150
		TG	-0.0002 (-0.78)	0.96204*** (29.93)								-0,0000 (-0.05)	0.7129

*[1] The significance at a two sided significance level of 10%, 5% and 1% is denoted with *, **, ***, respectively.*

Table IV: OLS Regression Models - Part II

Panel B: The Control Group: The Auditor's Non-Clients

Event date	Event name		Estimated OLS coefficients										Adj. R^2
			α	β	t	$t+1$	$t+2$	$t+3$	$t+4$	$t+5$	$[t, t+5]$	$[t, t+30]$	
26-01-2012	EY-I	CG	0.0000 (0.01)	1.0403*** (92.17)	0.0042 (1.63)	-0.0012 (-0.48)	0.0006 (0.23)	-0.0005 (-0.21)	0.0005 (0.18)	0.0021 (0.81)			0,9601
		CG	0.0000 (0.01)	1.0406*** (93.20)							0.0009 (0.88)		0,9603
		CG	-0.0000 (-0.17)	1.0407*** (93.27)								0.0005 (1.00)	0,9603
23-02-2012	Deloitte-I	CG	-0.0000 (-0.16)	1.0945*** (70.33)	0.0046 (1.26)	0.0021 (0.60)	0.0009 (0.26)	-0.0005 (-0.15)	0.0077** (2.14)	-0.0060* (-1.66)			0.9322
		CG	-0.0000 (-0.16)	1.0930*** (70.00)							0.0015 (0.99)		0.9316
		CG	-0.0000 (-0.14)	1.0931*** (69.92)								0.0002 (0.37)	0.9314
02-10-2012	EY-II	CG	-0.0001 (-0.34)	1.0640*** (91.08)	0.0037 (1.31)	0.0041 (1.44)	-0.0008 (-0.30)	0.0025 (0.89)	0.0072** (2.55)	-0.0019 (-0.66)			0.9585
		CG	-0.0001 (-0.34)	1.0638*** (90.94)							0.0024** (2.12)		0.9583
		CG	-0.0000 (-0.17)	1.0641*** (90.41)								0.0002 (0.33)	0.9578
03-06-2013	KPMG-I	CG	-0.0000 (-0.17)	1.0761*** (38.85)	-0.0019 (-0.43)	0.0007 (0.17)	0.0021 (0.49)	0.0037 (0.86)	0.0031 (0.71)	0.0053 (1.24)			0.8086
		CG	-0.0000 (-0.17)	1.0766*** (39.25)							0.0022 (1.24)		0.8104
		CG	-0.0000 (-0.02)	1.0751*** (39.14)								0.0000 (0.04)	0.8096
19-08-2013	KMPG-II	CG	0.0001 (0.54)	1.0786*** (45.63)	0.0009 (0.26)	-0.0011 (-0.33)	0.0014 (0.42)	0.0003 (0.08)	-0.0043 (-1.29)	0.0011 (0.32)			0.8529
		CG	0.0001 (0.54)	1.0775*** (45.92)							-0.0003 (-0.22)		0.8541
		CG	0.0001 (0.46)	1.0776*** (45.93)								0.0001 (0.13)	0.8541
30-12-2013	KPMG-III	CG	0.0003 (1.63)	1.0412*** (43.28)	0.0005 (0.18)	-0.0008 (-0.28)	-0.0003 (-0.09)	0.0033 (1.15)	0.0024 (0.82)	0.0013 (0.45)			0.8389
		CG	0.0003 (1.64)	1.0405*** (43.57)							0.0011 (0.91)		0.8405
		CG	0.0002 (1.16)	1.0412*** (43.76)								0.0010* (1.86)	0.8417

[1] *The significance at a two sided significance level of 10%, 5% and 1% is denoted with *, **, ***, respectively.*

V.I.I. THE TREATMENT GROUPS

Based on results depicted in table III and table IV, the cases of EY-I and DELOITTE-I yield significantly negative ARs on day t. Namely, these cases yield ARs of -2.12% (*t*=-2.43) and -0.86% (*t*=-1.76) on the event days respectively. Furthermore, KPMG-I provides a negative AR of -0.51% (*t*=-1.32) which is only significantly negative in case of a one sided significance level. However, none of the other events, i.e. EY-II, KPMG-II and KPMG-III, suggest that the publication leads to any significant abnormal market reaction on the publication date.

During the days following publication day t, only DELOITTE-I provides significant negative ARs. After insignificant ARs on the two days directly following the event date, the TG yields negative ARs of -0.85% (*t*=-1.74) and -0.97% (*t*=-1.98) on day *t*+3 and *t*+4 respectively. None of the other events have such significant market reactions that could support the hypothesis.[39] The remainder of the ARs are not significantly different from zero (0).

The CARs, i.e. the aggregation of the ARs during the different event windows, differ per event. Again, the case of DELOITTE-I provides the only significant negative CAR during the event period [t, t+5]. The overall impact on the TG's portfolio during this period is -3.67% (*t*=-3.05), suggesting that the news publication negatively affects the auditor's clients. Furthermore, the case of EY-I provides significant negative (C)ARs during the first few days after the publication, but this is no longer the case as of day t+3. The CAR of this event's TG is -3.20% (*t*=-2.12) during the period [t, t+2], which could support the research hypothesis. The CARs of both KPMG-I and KPMG-II during the event window [t, t+5] are negative, but not significant at any conventional significance level: -0.99% (*t*=-1.05) and -0.32% (*t*=-0.33) respectively. Furthermore, the CARs during the period [t, t+5] of EY-I, EY-II and KPMG-III are all positive but not statistically significant.

Except for the cases of EY-II and KPMG-II, the realized stock returns during the longest event window [t, t+30] are lower than expected in all cases. In short, in the 31-day period following the event, the TG yields negative CARs for almost all events. However, only in case of KPMG-I a significant market reaction is detected. The CAR of -4.31% (*t*=-2.02) in case of KPMG-I is significantly negative during the event window [t t+30]. As depicted in table III, none of the other CARs are significant at conventional significance levels.

Hence, the results are mixed per event. As such, the results do not consistently suggest that the TGs experience significant negative (C)ARs following the publication date. Based on the analysis of the (C)ARs, the hypothesis that the TG experiences negative (C)ARs following the publication of the news regarding their external auditor, cannot be confirmed consistently. However, the significant cases of EY-I, DELOITTE-I and – to a lesser extent – KPMG-I seem to be comparable with the findings of Firth (1990).[40] Firth detects a significant average AR of -1.1% during the week of publication, whereas the significance of the ARs vanishes after that particular week. All in all, the individual (C)AR analyses suggests that there is some (partial) evidence in favor of the hypothesis that auditor's clients experience negative abnormal

[39] In contrary, EY-I yields a positive market reaction (AR: +3.02% (*t*=3.47)).
[40] Note, however, that Firth's data of (abnormal) stock returns are on weekly basis, whereas the present study makes use of daily data.

returns directly after the publication of the enforcement initiatives. But in case a market reaction is actually detected, it seems to disappear after publication day t.

V.I.II. THE CONTROL GROUPS

As depicted in table III, the majority of the ARs of the CG on the day of publication is not significant. The only exception is EY-I, which provides a significantly positive AR of +0.44% (t=1.82). Hence, except EY-I, the ARs on the event date are predominantly positive, but not significant at any conventional significance level.

In the days following publication day t, a few significant ARs are detected. These significant ARs are observed for the DELOITTE-II and EY-II cases. Whereas the first experience a significant positive AR of +0.83% (t=2.59) on the fifth day since the market received the news, the latter experiences a daily AR of +0.71% (t=2.44) on that day. Note, however, that the positive market reaction in case of DELOITTE-II is largely offset with a negative AR of -0.55% (t=-1.71) at day t+4. Hence, these cases suggest that the event leads to neither a significantly positive market reaction nor a significantly negative market reaction during the days following the event. However, the ARs at least indicate that the CGs experience nonnegative ARs following the event.

As depicted in table III, the vast majority of the CARs during both event windows does not suggest any consistent market reaction. A potential spillover effect, as suggested in former studies, is not detected.

V.II. THE DIFFERENCES BETWEEN (C)ARS

Whereas the previous subsection focuses on each portfolio's (C)ARs separately, this subsection analyzes the differences between the portfolios following the events. As suggested by the research hypothesis, it is expected that the TG experiences lower ARs than its CG as counterpart. In order to assess this hypothesis, several parametric and nonparametric tests are performed. This subsection's remainder will centralize the outcomes of the unpaired t-test and the MWU test. However, any noteworthy differences between these results and the results of the paired t-tests and WSR tests will be discussed.

As mentioned before, the essence of the unpaired t-test is to detect whether the average difference between the ARs of the TG and CG are significantly different from zero (0). That is, it tests whether both the TG and CG are derived from an identical distributed population. Since the ARs experienced following the event date t are the central subjects of research, the tests are performed on the average differences during the event windows. Since the research question implies that the TG performs worse than the CG during the event windows, a one sided probability value is appropriate. The results of these tests are depicted in table V and discussed in this subsection's remainder.

Table V: The Analysis on Differences

Panel A (Panel B) depicts the test statistics of the parametric tests (nonparametric tests) on differences between average ARs during the event windows. For each test statistic, its particular t- or z-test statistic is denoted between brackets. Potential significance can be interpreted as described in footnote 1.

Panel A: The Parametric Analysis

Event name	Paired T-test		Unpaired T-test	
	[t,t+5] S	[t,t+30] S	[t,t+5] S	[t,t+30] S
EY-I	+ 0.06% (0.07)	+ 0.15% (0.43)	+ 0.06% (0.08)	+ 0.15% (0.43)
DELOITTE-I	+ 0.76% ** (2.60)	+ 0.08% (0.56)	+ 0.76% *** (3.32)	+ 0.08% (0.58)
EY-II	+ 0.20% (0.69)	- 0.15% (-0.95)	+ 0.20% (0.85)	- 0.15% (-1.02)
KPMG-I	+ 0.39% ** (3.22)	+ 0.16% *** (2.50)	+ 0.39% ** (2.36)	+ 0.16% ** (2.00)
KPMG-II	+ 0.03% (0.27)	-0.02% (-0.37)	+ 0.03% (0.24)	-0.02% (-0.34)
KMPG-III	- 0.13% (- 1.40)	+ 0.09% * (1.33)	- 0.13% (-1.35)	+ 0.09% (1.28)

Panel B: The Nonparametric Analysis

Event name	WSR test		MWU test	
	[t,t+5] S	[t,t+30] S	[t,t+5] S	[t,t+30] S
EY-I	-9 (-0.31)	-184 (-1.25)	14 (-0.64)	321 ** (-2.25)
DELOITTE-I	-1 ** (-1.99)	-225 (-0.45)	3 *** (-2.40)	465 (-0.22)
EY-II	-9 (-0.31)	+218 (-0.59)	14 (-0.64)	458 (0.32)
KPMG-I	-0 ** (-2.20)	-133 ** (-2.25)	5 ** (-2.08)	350 ** (-1.84)
KPMG-II	+9 (-0.31)	+223 (-0.49)	17 (0.16)	452 (0.40)
KMPG-III	+5 (-1.15)	-187 (-1.20)	10 (1.28)	364 * (-1.64)

*[1] The significance at a one sided significance level of 10%, 5% and 1% is denoted with *, **, ***, respectively.*

In line with the earlier discussed (C)AR analyses, the results of the unpaired t-tests differ between the individual events. The results indicate that the test's null hypothesis, holding that the average difference between the portfolios is not significantly different from zero (0), is rejected in the cases of DELOITTE-I and KPMG-I during the shorter event window [t, t+5]. The significant positive average differences (+0.76% (t=3.32) and +0.39% (t=2.36) respectively) suggest that the CG performs significantly better than the TG in these cases.[41] The other cases do not yield statistically significant differences in average ARs of their TG and CG. Hence, these cases do not support our research hypothesis. The results of the paired t-tests during the 6-day event window lead to comparable conclusions. Moreover, the nonparametric analyses, comprising the results of both the MWU and WSR tests, provide similar results during the short time horizon. Again, the test results in case of DELOITTE-I and KPMG-I indicate that the TGs perform significantly worse than their counterparts following the events,[42] and hence confirm the results from the t-tests.

For the extended time horizon, i.e. the event window covering the period [t, t+30], the unpaired t-tests only provide significant results for the KPMG-I case. The TG's ARs in this case are on average 0.16% lower than these of the CG, indicating that the daily ARs of the KPMG-I's clients are significantly lower than the ARs of the listed companies which appointed a different auditor for the event period. None of the other cases yield significant results. The paired t-test is in many respects comparable with the unpaired variant, since the KPMG-I case has significant negative results. In line with the unpaired t-test, the results for the EY-I and DELOITTE-I cases are positive but not statistically significant. However, some difference is found in case of KPMG-III. The outcomes in this case suggests that the TG performs significantly worse than the CG at a one sided 10% significance level by using the paired t-test.

The inferences that can be drawn from the nonparametric analyses on the 31-day event window are different from the analyses discussed above. Whereas the outcomes of the parametric analysis for KPMG-I and KPMG-III (for the paired t-test) cannot falsify the research hypothesis, the MWU test results indicate that the TG performs significantly worse than the CG in case of EY-I, KPMG-I and KPMG-III. The sum of ranks of the TGs, based on absolute differences, are significantly lower for these cases than the sum of ranks of their respective counterparts. The paired variant of the MWU test, i.e. the WSR test, only holds such outcomes for KPMG-I. The MWU test analyses provide some evidence that supports the hypothesis for the EY-I, KPMG-I and KPMG-III cases. However, the outcomes of the WSR tests, as included in table V, suggest that the TG performs significantly worse than the CG in case of KPMG-I only.

Hence, the analysis on differences between the portfolios does not bring the consistent conclusion that the TG performs worse than the CG. The test results of DELOITTE-I suggest that the average difference during the shorter time horizon is significantly positive, indicating that the CG outperforms the portfolio comprising DELOITTE-I's clients. The results of KPMG-I are more or less comparable to those of DELOITTE-I's. Moreover, the tests results of KPMG-I suggests that the TG performs significantly worse than the CG during the 31-day event window. As depicted in table V, the MWU tests in the EY-I and KPMG-

[41] Testing against either the one sided alternative hypothesis or the two sided alternative hypothesis does not lead to significantly different conclusions in case of the unpaired t-test.

[42] According to the '-'-sign of a particular test statistic, the sum of negative differences represents the test statistic. That is, the sum of positive differences sum is higher than the sum of negative differences. Based on the way these differences are conducted, i.e. $D_x = AR_{CG,x} - AR_{TG,x}$., these negative test statistics implies that the results are in line with the main research hypothesis.

III cases provide significant test results on the 31-day event window as well. The results for the other cases are not statistically significant.

V.III. THE ANALYSIS PER EVENT

The former two subsections discussed the test results. This subsection aggregates all the test results into one conclusion per event. The results are summarized in table VI below. The results are discussed in the remainder of this subsection.

Table VI: Summary of the Results

This table summarizes the results of all performed parametric and nonparametric tests. Instead of distinguishing the results between the different kind of tests, a distinction is made between the impact of the events on the short- (ST), medium- (MT) and the long-term (LT).

Event date	Cited name	Fine (EUR)		ST: [t]		MT: [t, t+5]		LT: [t, t+30]	
				TG	CG	TG	CG	TG	CG
Jan 26, 2012	EY-I	*54,450.-*	(C)Ars	-2.12%***	0.44%*	0.46%	0.79%	-2.72%	2.03%
			Concl.	Neg. effect	Pos. effect	No effect	No effect	No effect	No effect
			Diff.	N/A		+ 0.06%		+ 0.15%	(**)
			Concl.	N/A		No diff.		Pos. diff.	
Feb 23, 2012	DELOITTE-I	*54,450.-*	(C)Ars	-0.86%**	0.37%	-3.67%***	0.89%	-1.31%	1.15%
			Concl.	Neg. effect	No effect	Neg. effect	No effect	No effect	No effect
			Diff.	N/A		+ 0.76%	**(*)	+ 0.08%	
			Concl.	N/A		Pos. diff.		No diff.	
Oct 02, 2012	EY-II	*217,810.-*	(C)Ars	0.54%	0.36%	0.32%	1.53%**	5.46%	0.69%
			Concl.	No effect	No effect	No effect	Pos. effect	No effect	No effect
			Diff.	N/A		+ 0.20%		- 0.15%	
			Concl.	N/A		No diff.		No diff.	
Jun 03, 2013	KPMG-I	*881,250.-*	(C)Ars	-0.51%*	-0.18%	-0.99%	1.37%	-4.31%**	0.64%
			Concl.	Neg. effect	No effect	No effect	No effect	Neg. effect	No effect
			Diff.	N/A		+ 0.39%**		+ 0.16%	**(*)
			Concl.	N/A		Pos. diff.		Pos. diff.	
Aug 19, 2013	KMPG-II	*N/A*	(C)Ars	0.16%	0.13%	-0.32%	-0.14%	1.17%	0.51%
			Concl.	No effect	No effect	No effect	No effect	No effect	No effect
			Diff.	N/A		+ 0.03%		- 0.02%	
			Concl.	N/A		No diff.		No diff.	
Dec 30, 2013	KPMG-III	*7,000,000.-*	(C)Ars	0.54%	0.06%	1.41%	0.61%	-0.03%	2.83%
			Concl.	No effect	No effect	No effect	No effect	No effect	No effect
			Diff.	N/A		- 0.13%		+ 0.09%	(*)
			Concl.	N/A		No diff.		Pos. diff.	

*[1] The significance at a one sided significance level is used for the treatment group, and significance of 10%, 5% and 1% is denoted with *, **, ***, respectively. The significance at a two sided significance level is used for the control group, and significance of 10%, 5% and 1% is denoted with *, **, ***, respectively.*

EY-I

On January 26, 2012, the Dutch AFM makes use of its authority to publicly 'name and shame' an audit organization for the first time (Piersma, 2012a). The AFM decided to impose an administrative sanction (EUR 54,450) against EY for the lack of compliance with audit standards during the years 2007 and 2008. After assessing the TG's (and CG's) stock performance, a direct negative abnormal market reaction of -2.12% is detected for EY's clients. The conclusions based on the OLS regression model are largely in line with this. Moreover, the results indicate that the TG experiences lower (C)ARs than its CG during the 31-day event window.

The results of the parametric *t*-tests do not suggest that, on average, the CG significantly outperforms the TG during each of the event windows. The results of the nonparametric tests are not different from their parametric equivalents. The MWU test on the longest even window, however, suggests that the TG performs significantly worse than the CG.

Hence, the assessment of the EY-I case leads to the overall conclusion that a significant market reaction is detected on the publication day only, since the TG (CG) experiences a negative AR (positive AR). The MWU test during the 31-day event window only yields significant results that are in line with the research hypothesis. Since the remainder of the results (on the 31-day event window) does not suggest that the TG performs worse than its counterpart, consistent statistical evidence for a negative market reaction on the long term is not found.

DELOITTE-I

The case of DELOITTE-I concerns another administrative sanction (EUR 54,450), imposed by the AFM on February 23, 2012, because of Deloitte's non-compliance with audit standards during the year 2008. Therefore, the sanction is comparable to that of EY-I (Piersma, 2012b). In line with the results for EY-I, the TG's AR is significantly negative on the day of publication. The TG experiences a negative AR of -0.86%, whereas the CG does not experience any significant market reaction. Furthermore, the TG faces significantly negative ARs on day t+3 and t+4, which causes a negative CAR of -3.67% during the 6-day event window for the TG. For the longer event window, no such evidence can be obtained. As for the EY-I case, the OLS regression model leads to comparable results and similar conclusions.

Mainly caused by the significantly negative TG's ARs during the first days following the event, the parametric analyses suggest that the TG performs significantly worse than the CG during the short event window. The TG's ARs are on average 0.76% (t=3.32, according to the unpaired t-test) lower than those of the CG. Moreover, the nonparametric tests produce the same conclusion: according to both the signs and ranks of the experienced ARs, the TG experiences lower ARs during the shorter event window. This suggests that DELOITTE-I has a negative impact on the stock returns of its TG. However, the significance vanishes in case of the longer event window.

In short, DELOITTE-I brings the conclusion that a direct negative market reaction for the TG is detected on the day of publication and this effect continues for the 6-day short term event window. On the long term no evidence for such effect can be detected, indicating that the significant effects have (partly) reversed.

EY-II

The second administrative sanction has been imposed in 2012 on EY, i.e. EY-II, and amounts to a EUR 217,810 fee. Hereby it exceeds the amounts of the other earlier sanctions. The AFM-sanction was mainly based on insufficient performance of the function as a compliance officer as legally obliged for EY.[43] The results do not detect any (negative) market reaction for the TG. The CG, however, faces a positive CAR of 1.53% during the 6-day event window, whereas the TG's CAR during that event window is not significant. The CARs during the 31-day event window do not provide any significant results. The analyses on differences do not yield significant test results as well. Notably, the test results concerning the event window [t, t+30] hint that the average daily AR of the CG is 0.15% lower than the average of the TG. However, we note that this outcome is not significant.

The publication of EY-II does not lead to significantly negative (C)ARs for the TG. Hence, an adverse market reaction for the TG is not observed. Therefore, the EY-II case does not support the research hypothesis. The CG experiences a significantly positive CAR during the 5-day event window. During the 31-day event period, however, such positive market reaction is no longer observed.

KPMG-I

The remainder of the test results discussion focuses on the KPMG cases. These KPMG-cases are discussed in chronological order, starting with KPMG-I, representing the highest administrative sanction.[44] KPMG-I is a combination of two administrative fines (total EUR 881,250), because of an insufficient evaluation of the quality of the audits performed. Although the AFM and Het Financieele Dagblad (2013) both report that the fines are imposed on February 21, 2011 and definitive – after court decision – on May 30, 2013. The fines are published, and thus publicly known to all investors, on June 3, 2013.

The results of the (C)AR analyses suggest that the TG experiences a negative market reaction after the publication day t with a negative AR of -0.51%, whereas the CG does not experience any significant daily ARs. Although the CARs during the 6-day event window do not indicate any market reaction for the TG, the test outcome during the longer event windows [t, t+30] indicate that the TG experiences a negative CAR of -4.31%. This is checked against the results of the OLS regression model, which confirm this outcome.

This (relatively) negative stock performance of the TG is checked for robustness with the analyses on differences. These results are in line with those from the (C)AR analyses discussed before. The analyses on differences suggest that the TG's ARs are significantly worse during both event windows. The average AR of the TG is on average 0.39% lower than the CG during the 6-day event window, whereas this difference is decreased to 0.16% for the 31-day event window. The results are significant in both the parametric and nonparametric tests. All in all, the case of KPMG-I suggests that the publication of KPMG-I results in an adverse market reaction for the TG.

[43] A compliance officer is an internal supervisor that should control the quality and compliance of internal regulation. EY is obliged to appoint such a compliance officer according to art. 23 of the Audit Firms Supervision Decree (Bta).
[44] Note, KPMG-I is the highest fine that is imposed by the AFM. KPMG-III concerns an transaction between the auditor (i.e. KPMG) and the Dutch public prosecutor.

KPMG-II

The case of KPMG-II considers a different type of sanction. This event concerns the publication of the *Accountantskamer*'s decision regarding non-compliance with professional rules in case of Stichting Vestia. The results of KPMG-II do not provide any consistent evidence that the TG performs worse than the CG during the period following the publication. The OLS regression models do not provide evidence for a market reaction either. The results indicate that increasing the length of the event windows does not lead to any difference in conclusions. This could be due to the fact that the court's decision sees on a (former) employee of KPMG and not on the (entire or significant part of the) organization of KPMG; the judgment blames that the specific individual auditor responsible for the particular audit provided an unqualified audit opinion, whereas the court concludes that he was not allowed to provide such audit opinion based on the audit performed.

The tests on differences between the portfolios ARs do not support the research hypothesis either. The (non)parametric analyses imply that the TG does not experience significantly worse ARs following publication than its CG. Hence, both the individual (C)AR analysis as well as the tests on differences suggest that the TG does not perform (significantly) worse than its counterpart.

KPMG-III

KPMG-III considers a transaction between KPMG and the Dutch public prosecutor. The transaction involved an amount of EUR 7,000,000, consisting of a EUR 3,500,000 fine and a EUR 3,5000,000 compensation. This transaction closes the fraud case at Ballast Nedam that has been identified. Since none of the results of the (C)AR analyses are significant, any market reaction following the event is not detected. Again, the OLS regression models lead to a similar conclusion.

According to the analyses on differences, the results are mixed. The inferences based on the parametric t-tests suggest that the TG does not perform significantly worse than its counterpart during the 31-day event window. The MWU test, however, provides some evidence that the sum of ranks of the TG is significantly lower than the sum of ranks of the CG. However, this conclusion is not consistent with other tests. The analysis on differences during the shorter event window does not lead to any significant results.

All in all, a negative market reaction for the TG is not detected. Although the MWU test outcome suggest that the CG outperforms the TG during the 31-day event window, the results do not consistently confirm the research hypothesis. The absence of any consistent market reaction could be caused, because the transaction is merely related to KPMG-I's client, i.e. Ballast Nedam. Consistent with Callen and Morel's (2002) and Chaney and Philipich's (2002) findings, events directly related to the auditor seem to have a larger (negative) impact on client's stock returns than events related to the client involved.

VI. CONCLUSION

Theoretically, auditors should serve as gatekeepers for the company's investors by auditing the financial statements. However, various accounting scandals remind us that theory and practice are not always aligned. These scandals suggest that auditors may also have certain incentives to reduce their audit efforts. Considering the vast amount of public exposure on these cases, accounting scandals could possibly impair auditor reputation for delivering credible audit services.

Former studies already found some evidence, albeit mixed, that investors take into account auditor reputation while making investment decisions. However, these studies were primarily US focused and considered major accounting scandals, like Enron. The present study's approach is comparable to earlier studies, as it assesses the impact of (impaired) auditor reputation on their clients stock performance. However, our study focuses on the Dutch audit market specifically, an area which has remained untouched, and hereby contributes to our empirical knowledge.

We identified six events, that could serve as proxies for impaired auditor reputation. The events vary from news publications on administrative sanctions to the transaction of KPMG with the Dutch public prosecutor's of EUR 7,000,000 in the Ballast Nedam case.

In line with earlier studies, our results are mixed. A direct market reaction on the day of publication is detected in case of EY-I, DELOITTE-I and KPMG-I. In these cases, the clients experience significantly negative abnormal returns of -2.12%, -0.86% and -0.51% at conventional significance levels respectively. Although the auditor's clients do not experience a consistent market reaction following each of the events, the study's results suggest that the auditor's clients perform worse than their control group. Hence, the consequences of stressed auditor reputation seem to spill over onto the auditor's clients. Moreover, the results suggest that the stock returns of auditor's non-clients seem to be unaffected. Therefore, a potential spillover effect onto the rest of the market is not detected.

This thesis provides some interesting insights on the effect of impairment of auditor reputation on the financial performance of their audit clients. By our findings, suggesting a negative relationship exist between impairment of auditors and the stock returns of their other audit clients, it appears that the appointment of a particular auditor may represent a certain value for audited companies. Furthermore, such relationship does not exist for the auditor's non-clients. It also hints that a market effect exists that reaches beyond the direct monetary charge imposed on auditors and hereby relevant supervisors have a valuable – and as such effective – punitive instrument at their disposal. Although previous studies suggest that the reputation impairment of one particular auditor can imperil the entire market for auditing, this study does not detect a spillover effect onto the stock returns of non-clients.

We note that the availability of relevant events was limited for this study. We would therefore suggest future empirical studies to be conducted when the number of events increases. As the study is not geographically bound, the focus may also be extended to other countries. In addition, the future research can replicate this study to measure the effect on listed bonds.

VII. REFERENCES

Ball, R., & Brown, P. (1968). An Empirical Evaluation of Accounting Income Numbers. *Journal of Accounting Research, Vol. 6, No. 2* , pp. 159-78.

Barton, J. (2005). Who Cares about Auditor Reputation? *Contemporary Accounting Research, Vol. 22, No. 3* , pp. 549-86.

Binder, J. J. (1998). The Event Study Methodology Since 1969. *Review of Quantitative Finance and Accounting, Vol. 11* , pp. 111-37.

Bittlingmayer, G., & Hazlett, T. W. (2000). DOS Kapital: Has antitrust action against Microsoft created value in the computer industry? *Journal of Financial Economics, Vol. 55* , pp. 329-59.

Blume, M. E. (1971). On the Assessment of Risk. *The Journal of Finance, Vol. 26, No. 1* , pp. 1-10.

Boehmer, E., Musumeci, J., & Poulsen, A. B. (1991, October). Event-study methodology under conditions of event-induced variance. *Journal of Financial Economics, Vol. 30, No. 2* , pp. 253-72.

Brown, S. J., & Warner, J. B. (1980). Measuring Security Price Performance. *Journal of Financial Economics, Vol. 8* , pp. 205-58.

Brown, S. J., & Warner, J. B. (1985). Using Daily Stock Returns: The case of Event Studies. *Journal of Financial Economics, Vol. 14* , pp. 3-31.

Callen, J. L., & Morel, M. (2002). *The Enron-Andersen Debacle: Do Equity Markets React to Auditor Reputation?* EFMA 2003 Helsinki Meetings.

Chaney, P. K., & Philipich, K. L. (2002). Shredded Reputation: The Cost of Audit Failure. *Journal of Accounting Research, Vol. 40, No. 4* , pp. 1221-45.

Coffee Jr., J. C. (2002). *Understanding Enron: It's About the Gatekeepers, Stupid.* Columbia Law & Economics Working Paper No. 207.

Cowan, A. R. (1992). Nonparametric Event Study Tests. *Review of Quantitative Finance and Accounting, Vol. 2* , pp. 343-58.

DeAngelo, L. E. (1981). Auditor Size and Audit Quality. *Journal of Accounting and Economics, Vol. 3* , pp. 183-99.

Doogar, R., Sougiannis, T., & Xie, H. (2003). *The Impairment of Auditor Credibility: Stock Market Evidence from the Enron-Andersen Saga.*

Eckbo, B. (1983). Horizontal Mergers, Collusion, and Stockholder Wealth. *Journal of Financial Economics, Vol. 11* , pp. 241-74.

European Commission. (2010). *Audit Policy: Lessons From the Crisis.*

Fama, E. F. (1976). *Foundations of Finance: Portfolio Decisions and Securities Prices.* New York: Basic Books.

Fama, E. F., Fisher, L., Jensen, M. C., & Roll, R. (1969). The Adjustment of Stock Prices to New Information. *International Economic Review, Vol. 10* , pp. 1-21.

Firth, M. (1990). Auditor reputation: the impact of critical reports issued by government inspectors. *The RAND Journal of Economics, Vol. 21, No. 3* , pp. 374-87.

Fox, C. R., & Tversky, A. (1995). Ambiguity Aversion and Comparative Ignorance. *The Quarterly Journal of Economics, Vol. 110, No. 3* , pp. 585-603.

Gonedes, N. J. (1973). Evidence on the Information Content of Accounting Numbers: Accounting-Based and Market-Based Estimates of Systematic Risk. *Journal of Financial and Quantitative Analysis* , pp. 407-44.

Het Financieele Dagblad. (2012, October 3). Tweede boete in een jaar. *Het Financieele Dagblad* , p. 13.

Izan, H. Y. (1978). *An empirical analysis of the economic effecs of mandatory government audit requirements.* Dissertation, University of Chicago.

Jensen, M. C., & Meckling, W. H. (1976). Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure. *Journal of Financial Economics, Vol. 3, No. 4* , pp. 305-60.

Josev, T., Chan, H., & Faff, R. (2004). What's in a Name? Evidence on Corporate Name Changes from the Australian Capital Market. *Pacific Accounting Review, Vol. 16, No. 1* , pp. 57-75.

Kahneman, D., & Tversky, A. (1984). Choices, Values and Frames. *American Psychologist, Vol. 39, No. 4* , pp. 341-50.

Klein, B., & Leffler, K. (1981). The Role of Market Forces in Assuring Contractual Performance. *Journal of Political Economy, Vol. 89, No. 41* , pp. 615-41.

Kothari, S., & Warner, J. B. (2007). Econometrics of Event Studies. In E. Eckbo, *Handbook of Corporate Finance: Empirical Corporate Finance.* Amsterdam: Elsevier.

Levitt, A. (1998). *The numbers game. Remarks delivered at NYU Center for Law and Business.* New York University, September 28.

MacKinlay, A. G. (1997). Event Studies in Economics and Finance. *Journal of Economic Literature, Vol. 35, No. 1* , pp. 13-39.

Piersma, J. (2012, February 24). AFM geeft ook Deloitte boete. *Het Financieele Dagblad* , p. 9.

Piersma, J. (2012, January 27). AFM laat tanden zien in nieuw toezichtsdomein. *Het Financieele Dagblad* , p. 11.

Serra, A. P. (2002). *Event Study Tests: A brief survey.* Faculdade de Economia, Universidade Do Porto WP No. 117.

Sharpe, W. F. (1963). A Simplified Model for Portfolio Analysis. *Management Science* , pp. 277-93.

Siegel, S. (1957). Nonparametric Statistics. *The American Statistician, Vol. 11, No. 3* , pp. 13-9.

Skinner, D. J., & Srinivasan, S. (2012). Audit Quality and Auditor Reputation: Evidence from Japan. *The Accounting Review, Vol. 87, No. 5* , pp. 1737-65.

Soroka, S. N. (2006). Good News and Bad News: Asymmetric Responses to Economic Information. *The Journal of Politics, Vol. 68, No. 2* , pp. 372-85.

Tversky, A., & Kahneman, D. (1991). Loss Aversion in Riskless Choice: A Reference-Depedendent Model. *The Quarterly Journal of Economics, Vol. 107, No. 4* , pp. 1039-61.

Watts, R. L., & Zimmerman, J. L. (1983). Agency Problems, Auditing, and The Theory of the Firm: Some Evidence. *Journal of Law and Economics, Vol. 26, No. 3* , pp. 613-33.

Weber, J., Willenborg, M., & Zhang, J. (2008). Does Auditor Reputation Matter? The Case of KPMG Germany and ComROAD AG. *Journal of Accounting Research, Vol. 46, No. 4* , pp. 941-72.

Wit, J., Braakhuis, B., Haverkamp, M., Koser-Kaya, F., Neppérus, H., & Vermeij, R. (2010, May 10). *Verloren krediet.*

VIII. APPENDICES

Appendix A: The Construction of the Portfolios - Part I

This table includes the companies that are used in the portfolio construction per event. Sequentially, for each event, the full company name, its index as well as its external auditor is depicted. Moreover, TG, CG and CI indicates whether the company is included in the treatment group, control group or omitted since it is the relevant company involved, respectively.

EY-I				Deloitte-I				EY-II			
Company	Index	Auditor	TG / CG / CI	Company	Index	Auditor	TG / CG / CI	Company	Index	Auditor	TG / CG / CI
Aegon N.V.	AEX	EY	CG	Aegon N.V.	AEX	EY	CG	Aegon N.V.	AEX	EY	TG
Koninklijke Ahold N.V.	AEX	Deloitte	CG	Koninklijke Ahold N.V.	AEX	Deloitte	TG	Koninklijke Ahold N.V.	AEX	Deloitte	CG
Akzo Nobel N.V.	AEX	KPMG	CG	Akzo Nobel N.V.	AEX	KPMG	CG	Akzo Nobel N.V.	AEX	KPMG	CG
ArcelorMittal S.A.	AEX	Deloitte	CG	ArcelorMittal S.A.	AEX	Deloitte	TG	Aperam S.A.	AEX	Deloitte	CG
ASML Holding N.V.	AEX	Deloitte	CG	ASML Holding N.V.	AEX	Deloitte	TG	ArcelorMittal S.A.	AEX	Deloitte	CG
Boskalis	AEX	KPMG	CG	Boskalis	AEX	KPMG	CG	ASML Holding N.V.	AEX	Deloitte	CG
Corio N.V.	AEX	PWC	CG	Corio N.V.	AEX	PWC	CG	Boskalis	AEX	KPMG	CG
Koninklijke DSM N.V.	AEX	EY	CG	Koninklijke DSM N.V.	AEX	EY	CG	Corio N.V.	AEX	PWC	CG
Fugro N.V.	AEX	KPMG	CG	Fugro N.V.	AEX	KPMG	CG	Koninklijke DSM N.V.	AEX	EY	TG
Heineken N.V.	AEX	KPMG	CG	Heineken N.V.	AEX	KPMG	CG	Fugro N.V.	AEX	KPMG	CG
ING Groep N.V.	AEX	EY	CG	ING Groep N.V.	AEX	EY	CG	Heineken N.V.	AEX	KPMG	CG
Koninklijke KPN N.V.	AEX	PWC	CG	Koninklijke KPN N.V.	AEX	PWC	CG	ING Groep N.V.	AEX	EY	TG
Koninklijke Philips N.V.	AEX	KPMG	CG	Koninklijke Philips N.V.	AEX	KPMG	CG	Koninklijke KPN N.V.	AEX	PWC	CG
PostNL N.V.	AEX	PWC	CG	PostNL N.V.	AEX	PWC	CG	Koninklijke Philips N.V.	AEX	KPMG	CG
Randstad Holding N.V.	AEX	PWC	CG	Randstad Holding N.V.	AEX	PWC	CG	PostNL N.V.	AEX	PWC	CG
Reed Elsevier N.V.	AEX	Deloitte	CG	Reed Elsevier N.V.	AEX	Deloitte	TG	Randstad Holding N.V.	AEX	PWC	CG
Royal Dutch Shell Plc	AEX	PWC	CG	Royal Dutch Shell Plc	AEX	PWC	CG	Reed Elsevier N.V.	AEX	Deloitte	CG
SBM Offshore N.V.	AEX	KPMG	CG	SBM Offshore N.V.	AEX	KPMG	CG	Royal Dutch Shell Plc	AEX	PWC	CG
TomTom N.V.	AEX	Deloitte	CG	TomTom N.V.	AEX	Deloitte	TG	SBM Offshore N.V.	AEX	KPMG	CG
Unibail-Rodamco SE	AEX	Deloitte; EY	CG	Unibail-Rodamco SE	AEX	Deloitte; EY	CG	TNT Express N.V.	AEX	PWC	CG
Unilever N.V.	AEX	PWC	CG	Unilever N.V.	AEX	PWC	CG	Unibail-Rodamco SE	AEX	Deloitte; EY	CG
Wolters Kluwer N.V.	AEX	KPMG	CG	Wolters Kluwer N.V.	AEX	KPMG	CG	Unilever N.V.	AEX	PWC	CG
Aalberts Industries N.V.	AMX	PWC	CG	Aalberts Industries N.V.	AMX	PWC	CG	Wolters Kluwer N.V.	AEX	KPMG	CG
AMG Advanced Metallugical Group N.V.	AMX	EY	CG	AMG Advanced Metallugical Group N.V.	AMX	EY	CG	Aalberts Industries N.V.	AMX	PWC	CG
Arcadis N.V.	AMX	KPMG	CG	Arcadis N.V.	AMX	KPMG	CG	AMG Advanced Metallugical Group N.V.	AMX	EY	TG
ASM International N.V.	AMX	Deloitte	CG	ASM International N.V.	AMX	Deloitte	TG	Arcadis N.V.	AMX	KPMG	CG
Koninklijke BAM Groep N.V.	AMX	PWC	CG	Koninklijke BAM Groep N.V.	AMX	PWC	CG	ASM International N.V.	AMX	Deloitte	CG
BinckBank N.V.	AMX	EY	CG	BinckBank N.V.	AMX	EY	CG	Koninklijke BAM Groep N.V.	AMX	PWC	CG
Brunel International N.V.	AMX	Deloitte	CG	Brunel International N.V.	AMX	Deloitte	TG	BinckBank N.V.	AMX	EY	TG
Corbion N.V.	AMX	Deloitte	CG	Corbion N.V.	AMX	Deloitte	TG	Brunel International N.V.	AMX	Deloitte	CG
Delta Lloyd N.V.	AMX	EY	CG	Delta Lloyd N.V.	AMX	EY	CG	Corbion N.V.	AMX	Deloitte	CG
Eurocommercial Properties N.V.	AMX	EY	CG	Eurocommercial Properties N.V.	AMX	EY	CG	Delta Lloyd N.V.	AMX	EY	TG
Heijmans N.V.	AMX	KPMG	CG	Heijmans N.V.	AMX	KPMG	CG	Eurocommercial Properties N.V.	AMX	EY	TG
Royal Imtech N.V.	AMX	KPMG	CG	Royal Imtech N.V.	AMX	KPMG	CG	Heijmans N.V.	AMX	KPMG	CG
Logica PLC	AMX	PWC	CG	Logica PLC	AMX	PWC	CG	Mediq N.V.	AMX	KPMG	CG
Mediq N.V.	AMX	KPMG	CG	Mediq N.V.	AMX	KPMG	CG	NSI N.V.	AMX	KPMG	CG
Nutreco N.V.	AMX	KPMG	CG	Nutreco N.V.	AMX	KPMG	CG	Nutreco N.V.	AMX	KPMG	CG
Ordina N.V.	AMX	PWC	CG	Ordina N.V.	AMX	PWC	CG	Royal Imtech N.V.	AMX	KPMG	CG
Pharming Group N.V.	AMX	PWC	CG	Pharming Group N.V.	AMX	PWC	CG	SNS Reaal N.V.	AMX	KPMG	CG
SNS Reaal N.V.	AMX	KPMG	CG	SNS Reaal N.V.	AMX	KPMG	CG	Koninklijke Ten Cate N.V.	AMX	KPMG	CG
Koninklijke Ten Cate N.V.	AMX	KPMG	CG	Koninklijke Ten Cate N.V.	AMX	KPMG	CG	TomTom N.V.	AMX	Deloitte	CG
Unit4 N.V.	AMX	EY	CG	Unit4 N.V.	AMX	EY	CG	Unit4 N.V.	AMX	EY	TG
USG People N.V.	AMX	PWC	CG	USG People N.V.	AMX	PWC	CG	USG People N.V.	AMX	PWC	CG
VastNed Retail N.V.	AMX	Deloitte	CG	VastNed Retail N.V.	AMX	Deloitte	TG	VastNed Retail N.V.	AMX	Deloitte	CG
Koninklijke Vopak N.V.	AMX	PWC	CG	Koninklijke Vopak N.V.	AMX	PWC	CG	Koninklijke Vopak N.V.	AMX	PWC	CG
Wavin N.V.	AMX	Deloitte	CG	Wavin N.V.	AMX	Deloitte	TG	Wereldhave N.V.	AMX	PWC	CG
Wereldhave N.V.	AMX	PWC	CG	Wereldhave N.V.	AMX	PWC	CG	Koninklijke Wessanen N.V.	AMX	Deloitte	CG
Accell Group N.V.	AScX	Deloitte	CG	Accell Group N.V.	AScX	Deloitte	TG	Accell Group N.V.	AScX	Deloitte	CG
Amsterdam Commodities N.V.	AScX	BDO	CG	Amsterdam Commodities N.V.	AScX	BDO	CG	Amsterdam Commodities N.V.	AScX	BDO	CG
AMT Holding N.V.	AScX	PWC	CG	AMT Holding N.V.	AScX	PWC	CG	Ballast Nedam N.V.	AScX	KPMG	CG
Antonov PLC	AScX	EY	CG	Antonov PLC	AScX	EY	CG	BE Semiconductor Industries N.V.	AScX	KPMG	CG
Ballast Nedam N.V.	AScX	KPMG	CG	Ballast Nedam N.V.	AScX	KPMG	CG	Beter Bed Holding N.V.	AScX	EY	TG
BE Semiconductor Industries N.V.	AScX	KPMG	CG	BE Semiconductor Industries N.V.	AScX	KPMG	CG	Cryo-Save Group N.V.	AScX	KPMG	CG
Beter Bed Holding N.V.	AScX	EY	CG	Beter Bed Holding N.V.	AScX	EY	CG	DOCDATA N.V.	AScX	KPMG	CG
Dockwise Ltd.	AScX	KPMG	CG	Dockwise Ltd.	AScX	KPMG	CG	Dockwise Ltd.	AScX	KPMG	CG
Exact Holding N.V.	AScX	KPMG	CG	Exact Holding N.V.	AScX	KPMG	CG	Exact Holding N.V.	AScX	KPMG	CG
Fornix BioSciences N.V.	AScX	KPMG	CG	Fornix BioSciences N.V.	AScX	KPMG	CG	Grontmij N.V.	AScX	Deloitte	CG
Grontmij N.V.	AScX	Deloitte	CG	Grontmij N.V.	AScX	Deloitte	TG	ICT Automatisering N.V.	AScX	Deloitte	CG
Kardan N.V.	AScX	EY	CG	Kardan N.V.	AScX	EY	CG	Kardan N.V.	AScX	EY	TG
Kas Bank N.V.	AScX	KPMG	CG	Kas Bank N.V.	AScX	KPMG	CG	Kas Bank N.V.	AScX	KPMG	CG
Kendrion N.V.	AScX	KPMG	CG	Kendrion N.V.	AScX	KPMG	CG	Kendrion N.V.	AScX	KPMG	CG
LBi International N.V.	AScX	PWC	CG	LBi International N.V.	AScX	PWC	CG	LBi International N.V.	AScX	PWC	CG
Macintosh Retail Groep N.V.	AScX	EY	CG	Macintosh Retail Groep N.V.	AScX	EY	CG	Macintosh Retail Groep N.V.	AScX	EY	TG
N.V. Nederlandsche Apparatenfabriek 'Nedap'	AScX	KPMG	CG	N.V. Nederlandsche Apparatenfabriek 'Nedap'	AScX	KPMG	CG	N.V. Nederlandsche Apparatenfabriek 'Nedap'	AScX	KPMG	CG
NSI N.V.	AScX	KPMG	CG	NSI N.V.	AScX	KPMG	CG	Ordina N.V.	AScX	PWC	CG
Sligro Food Group N.V.	AScX	KPMG	CG	Sligro Food Group N.V.	AScX	KPMG	CG	Pharming Group N.V.	AScX	PWC	CG
Telegraaf Media Groep N.V.	AScX	Deloitte	CG	Telegraaf Media Groep N.V.	AScX	Deloitte	TG	RoodMicrotec N.V.	AScX	Mazars	CG
TKH Group N.V.	AScX	Deloitte	CG	TKH Group N.V.	AScX	Deloitte	TG	Sligro Food Group N.V.	AScX	KPMG	CG
Koninklijke Wessanen N.V.	AScX	Deloitte	CG	Koninklijke Wessanen N.V.	AScX	Deloitte	TG	TKH Group N.V.	AScX	Deloitte	CG
Xeikon N.V.	AScX	BDO	CG	Xeikon N.V.	AScX	BDO	CG	Xeikon N.V.	AScX	BDO	CG

[1] *Omitted companies are Air France KLM S.A. (CG), Aperam S.A. (CG), TNT Express N.V. (CG), Arseus N.V. (CG) and Qurius N.V. (CG) in case of EY-I; Air France KLM N.V. (CG), peram S.A. (TG), TNT Express N.V. (CG), Arseus N.V. (CG), Qurius N.V. (CG); and Air France KLM S.A. (CG), D.E. Master Blenders (CG), Ziggo N.V. (TG), Arseus N.V. (CG) and Qurius N.V. (CG) in case of EY-II.*

Appendix A: The Construction of the Portfolios - Part II

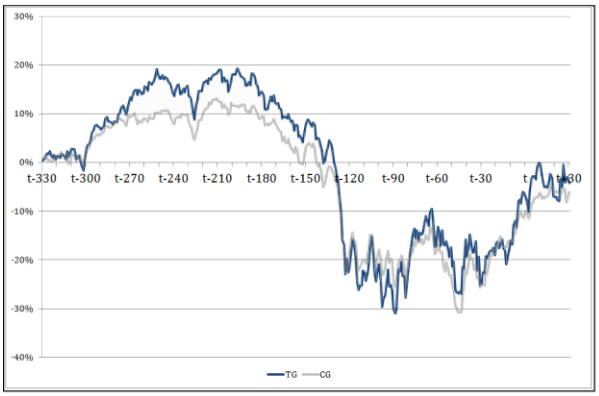
EY-I				Deloitte-I				EY-II			
Company	Index	Auditor	TG / CG / CI	Company	Index	Auditor	TG / CG / CI	Company	Index	Auditor	TG / CG / CI
Aegon N.V.	AEX	EY	CG	Aegon N.V.	AEX	EY	CG	Aegon N.V.	AEX	EY	CG
Koninklijke Ahold N.V.	AEX	PWC	CG	Koninklijke Ahold N.V.	AEX	PWC	CG	Koninklijke Ahold N.V.	AEX	PWC	CG
Akzo Nobel N.V.	AEX	KPMG	TG	Akzo Nobel N.V.	AEX	KPMG	TG	Akzo Nobel N.V.	AEX	KPMG	TG
Aperam S.A.	AEX	Deloitte	CG	ArcelorMittal S.A.	AEX	Deloitte	CG	ArcelorMittal S.A.	AEX	Deloitte	CG
ArcelorMittal S.A.	AEX	Deloitte	CG	ASML Holding N.V.	AEX	Deloitte	CG	ASML Holding N.V.	AEX	Deloitte	CG
ASML Holding N.V.	AEX	Deloitte	CG	Corio N.V.	AEX	PWC	CG	Corio N.V.	AEX	PWC	CG
Corio N.V.	AEX	PWC	CG	Koninklijke DSM N.V.	AEX	EY	CG	Koninklijke DSM N.V.	AEX	EY	CG
Koninklijke DSM N.V.	AEX	EY	CG	Fugro N.V.	AEX	KPMG	TG	Fugro N.V.	AEX	KPMG	TG
Fugro N.V.	AEX	KPMG	TG	Gemalto N.V.	AEX	PWC	CG	Gemalto N.V.	AEX	PWC	CG
Heineken N.V.	AEX	KPMG	TG	Heineken N.V.	AEX	KPMG	TG	Heineken N.V.	AEX	KPMG	TG
ING Groep N.V.	AEX	EY	CG	ING Groep N.V.	AEX	EY	CG	ING Groep N.V.	AEX	EY	CG
Koninklijke KPN N.V.	AEX	PWC	CG	Koninklijke KPN N.V.	AEX	PWC	CG	Koninklijke KPN N.V.	AEX	PWC	CG
Koninklijke Philips N.V.	AEX	KPMG	TG	Koninklijke Philips N.V.	AEX	KPMG	TG	Koninklijke Philips N.V.	AEX	KPMG	TG
PostNL N.V.	AEX	PWC	CG	PostNL N.V.	AEX	PWC	CG	PostNL N.V.	AEX	PWC	CG
Randstad Holding N.V.	AEX	PWC	CG	Randstad Holding N.V.	AEX	PWC	CG	Randstad Holding N.V.	AEX	PWC	CG
Reed Elsevier N.V.	AEX	Deloitte	CG	Reed Elsevier N.V.	AEX	Deloitte	CG	Reed Elsevier N.V.	AEX	Deloitte	CG
Royal Dutch Shell Plc	AEX	PWC	CG	Royal Dutch Shell Plc	AEX	PWC	CG	Royal Dutch Shell Plc	AEX	PWC	CG
Royal Imtech N.V.	AEX	KPMG	TG	Royal Imtech N.V.	AEX	KPMG	TG	Royal Imtech N.V.	AEX	KPMG	TG
SBM Offshore N.V.	AEX	KPMG	TG	SBM Offshore N.V.	AEX	KPMG	TG	SBM Offshore N.V.	AEX	KPMG	TG
TNT Express N.V.	AEX	PWC	CG	TNT Express N.V.	AEX	PWC	CG	TNT Express N.V.	AEX	PWC	CG
Unibail-Rodamco SE	AEX	Deloitte; EY	CG	Unibail-Rodamco SE	AEX	Deloitte; EY	CG	Unibail-Rodamco SE	AEX	Deloitte; EY	CG
Unilever N.V.	AEX	PWC	CG	Unilever N.V.	AEX	PWC	CG	Unilever N.V.	AEX	PWC	CG
Wolters Kluwer N.V.	AEX	KPMG	TG	Wolters Kluwer N.V.	AEX	KPMG	TG	Wolters Kluwer N.V.	AEX	KPMG	TG
Aalberts Industries N.V.	AMX	PWC	CG	Aalberts Industries N.V.	AMX	PWC	CG	Ziggo N.V.	AEX	EY	CG
AMG Advanced Metallugical Group N.V.	AMX	EY	CG	AMG Advanced Metallugical Group N.V.	AMX	EY	CG	Aalberts Industries N.V.	AMX	PWC	CG
Arcadis N.V.	AMX	KPMG	TG	Aperam S.A.	AMX	Deloitte	CG	AMG Advanced Metallugical Group N.V.	AMX	EY	CG
ASM International N.V.	AMX	Deloitte	CG	Arcadis N.V.	AMX	KPMG	TG	Aperam S.A.	AMX	Deloitte	CG
Koninklijke BAM Groep N.V.	AMX	PWC	CG	ASM International N.V.	AMX	Deloitte	CG	Arcadis N.V.	AMX	KPMG	TG
BinckBank N.V.	AMX	EY	CG	Koninklijke BAM Groep N.V.	AMX	PWC	CG	ASM International N.V.	AMX	Deloitte	CG
Brunel International N.V.	AMX	PWC	CG	BinckBank N.V.	AMX	EY	CG	Koninklijke BAM Groep N.V.	AMX	PWC	CG
Corbion N.V.	AMX	Deloitte	CG	Koninklijke Boskalis Westminster N.V.	AMX	KPMG	TG	BinckBank N.V.	AMX	EY	CG
Delta Lloyd N.V.	AMX	EY	CG	Brunel International N.V.	AMX	PWC	CG	Koninklijke Boskalis Westminster N.V.	AMX	KPMG	TG
Eurocommercial Properties N.V.	AMX	EY	CG	Corbion N.V.	AMX	Deloitte	CG	Brunel International N.V.	AMX	PWC	CG
Heijmans N.V.	AMX	KPMG	TG	Delta Lloyd N.V.	AMX	EY	CG	Corbion N.V.	AMX	Deloitte	CG
Koninklijke Koninklijke Boskalis Westminster N.V.	AMX	KPMG	TG	Eurocommercial Properties N.V.	AMX	EY	CG	Delta Lloyd N.V.	AMX	EY	CG
NSI N.V.	AMX	KPMG	TG	Heijmans N.V.	AMX	KPMG	TG	Eurocommercial Properties N.V.	AMX	EY	CG
Nutreco N.V.	AMX	KPMG	TG	NSI N.V.	AMX	KPMG	TG	Heijmans N.V.	AMX	KPMG	TG
Pharming Group N.V.	AMX	PWC	CG	Nutreco N.V.	AMX	KPMG	TG	NSI N.V.	AMX	KPMG	TG
Koninklijke Ten Cate N.V.	AMX	KPMG	TG	Pharming Group N.V.	AMX	PWC	CG	Nutreco N.V.	AMX	KPMG	TG
TKH Group N.V.	AMX	Deloitte	CG	Koninklijke Ten Cate N.V.	AMX	KPMG	CG	Pharming Group N.V.	AMX	PWC	CG
TomTom N.V.	AMX	Deloitte	CG	TKH Group N.V.	AMX	Deloitte	CG	Koninklijke Ten Cate N.V.	AMX	KPMG	TG
Unit4 N.V.	AMX	EY	CG	TomTom N.V.	AMX	Deloitte	CG	TKH Group N.V.	AMX	Deloitte	CG
USG People N.V.	AMX	PWC	CG	Unit4 N.V.	AMX	EY	CG	TomTom N.V.	AMX	Deloitte	CG
VastNed Retail N.V.	AMX	Deloitte	CG	USG People N.V.	AMX	PWC	CG	Unit4 N.V.	AMX	EY	CG
Koninklijke Vopak N.V.	AMX	PWC	CG	VastNed Retail N.V.	AMX	Deloitte	CG	USG People N.V.	AMX	PWC	CG
Wereldhave N.V.	AMX	PWC	CG	Koninklijke Vopak N.V.	AMX	PWC	CG	VastNed Retail N.V.	AMX	Deloitte	CG
Koninklijke Wessanen N.V.	AMX	Deloitte	CG	Wereldhave N.V.	AMX	PWC	CG	Koninklijke Vopak N.V.	AMX	PWC	CG
Accell Group N.V.	AScX	Deloitte	CG	Ziggo N.V.	AMX	EY	CG	Wereldhave N.V.	AMX	PWC	CG
Amsterdam Commodities N.V.	AScX	BDO	CG	Accell Group N.V.	AScX	Deloitte	CG	Accell Group N.V.	AScX	Deloitte	CG
AND International Publishers N.V.	AScX	Mazars	CG	Amsterdam Commodities N.V.	AScX	BDO	CG	Amsterdam Commodities N.V.	AScX	BDO	CG
Ballast Nedam N.V.	AScX	EY	CG	AND International Publishers N.V.	AScX	Mazars	CG	AND International Publishers N.V.	AScX	Mazars	CG
Batenburg Techniek	AScX	Mazars	CG	Ballast Nedam N.V.	AScX	EY	TG	Ballast Nedam N.V.	AScX	EY	CI
BE Semiconductor Industries N.V.	AScX	KPMG	TG	BE Semiconductor Industries N.V.	AScX	KPMG	CG	BE Semiconductor Industries N.V.	AScX	KPMG	TG
Beter Bed Holding N.V.	AScX	EY	CG	Beter Bed Holding N.V.	AScX	EY	CG	Beter Bed Holding N.V.	AScX	EY	CG
Koninklijke Brill N.V.	AScX	EY	CG	Koninklijke Brill N.V.	AScX	EY	CG	Koninklijke Brill N.V.	AScX	EY	CG
Crown Van Gelder N.V.	AScX	PWC	CG	Crown Van Gelder N.V.	AScX	PWC	TG	Crown Van Gelder N.V.	AScX	PWC	CG
Cryo-Save Group N.V.	AScX	KPMG	TG	Cryo-Save Group N.V.	AScX	KPMG	TG	Cryo-Save Group N.V.	AScX	KPMG	TG
DOCDATA N.V.	AScX	KPMG	TG	DOCDATA N.V.	AScX	KPMG	TG	DOCDATA N.V.	AScX	KPMG	TG
Grontmij N.V.	AScX	Deloitte	CG	Exact Holding N.V.	AScX	KPMG	CG	Exact Holding N.V.	AScX	KPMG	TG
Holland Colours N.V.	AScX	EY	CG	Grontmij N.V.	AScX	Deloitte	CG	Grontmij N.V.	AScX	Deloitte	CG
ICT Automatisering N.V.	AScX	Deloitte	CG	Holland Colours N.V.	AScX	EY	CG	Holland Colours N.V.	AScX	EY	CG
Kardan N.V.	AScX	EY	CG	ICT Automatisering N.V.	AScX	Deloitte	CG	ICT Automatisering N.V.	AScX	Deloitte	CG
Kas Bank N.V.	AScX	KPMG	TG	Kardan N.V.	AScX	EY	TG	Kardan N.V.	AScX	EY	CG
Kendrion N.V.	AScX	KPMG	TG	Kas Bank N.V.	AScX	KPMG	TG	Kas Bank N.V.	AScX	KPMG	TG
N.V. Nederlandsche Apparatenfabriek 'Nedap'	AScX	KPMG	TG	Kendrion N.V.	AScX	KPMG	TG	Kendrion N.V.	AScX	KPMG	TG
Ordina N.V.	AScX	PWC	CG	N.V. Nederlandsche Apparatenfabriek 'Nedap'	AScX	KPMG	CG	N.V. Nederlandsche Apparatenfabriek 'Nedap'	AScX	KPMG	TG
RoodMicrotec N.V.	AScX	Mazars	CG	Ordina N.V.	AScX	PWC	TG	Ordina N.V.	AScX	PWC	CG
Sligro Food Group N.V.	AScX	KPMG	TG	Sligro Food Group N.V.	AScX	KPMG	CG	Sligro Food Group N.V.	AScX	KPMG	TG
Value8 N.V.	AScX	Mazars PH	CG	Telegraaf Media Groep N.V.	AScX	Deloitte	CG	Telegraaf Media Groep N.V.	AScX	Deloitte	CG
Vivenda Media Groep N.V.	AScX	Baker Tilly Berk	CG	Value8 N.V.	AScX	Mazars PH	CG	Value8 N.V.	AScX	Mazars PH	CG
Xeikon N.V.	AScX	BDO	CG	Koninklijke Wessanen N.V.	AScX	Deloitte	CG	Koninklijke Wessanen N.V.	AScX	Deloitte	CG
				Xeikon N.V.	AScX	BDO		Xeikon N.V.	AScX	BDO	CG

[1] Omitted companies are Air France KLM S.A. (CG), D.E. Master Blenders (CG), Ziggo N.V. (CG) and Arseus N.V. (CG) in case of KPMG-I; Air France KLM N.V. (CG), D.E. Master Blender N.V. (CG) and Arseus N.V. (CG) in case of KPMG-II; and Air France KLM S.A. (CG), OCI N.V. (CG) and Arseus N.V. (TG) in case of KPMG-III.

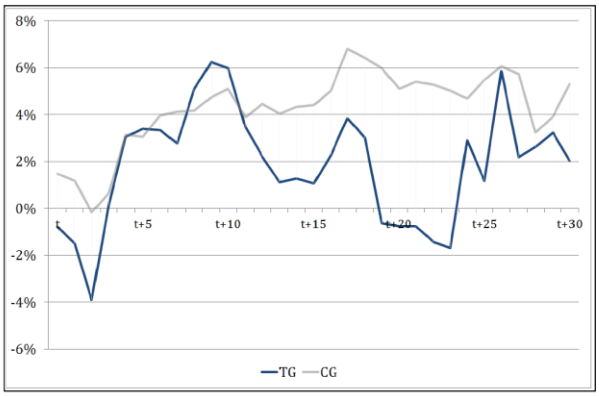
Appendix B: Cumulative Returns – Part I

The left figure depict graphical representations of the cumulative returns during both the estimation and the event period for each event separately. The right figures show a zoom on the event period specifically.

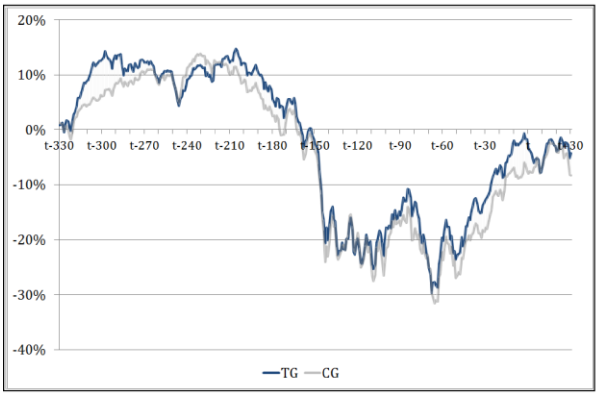
EY-I



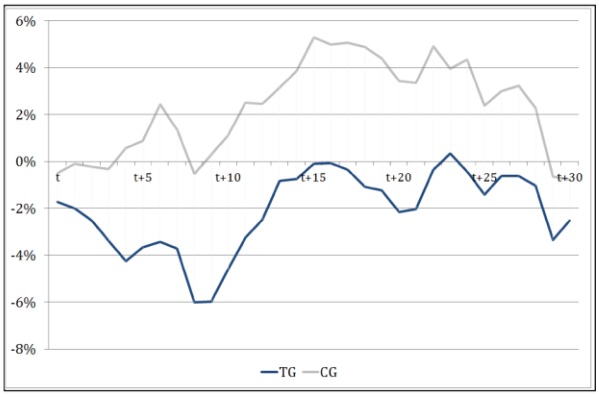
EY-I



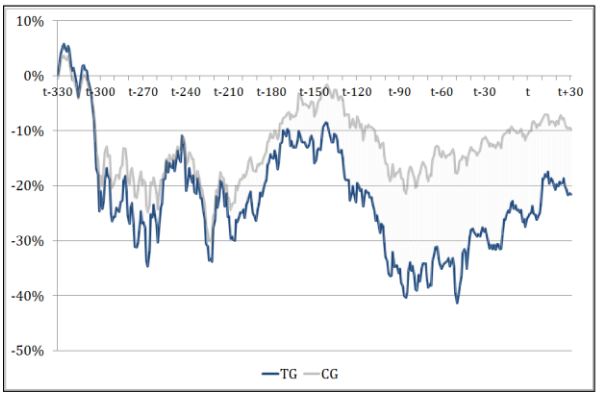
Deloitte-I



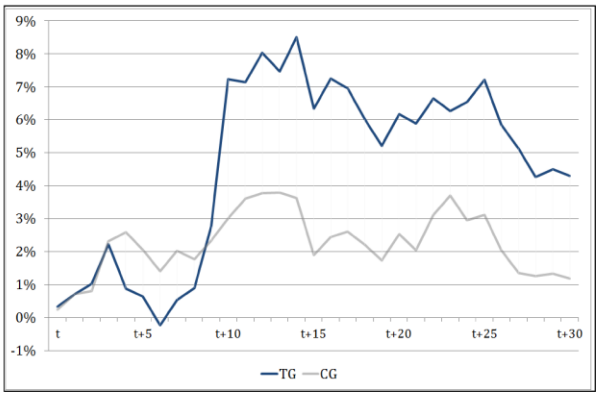
Deloitte-I



EY-II

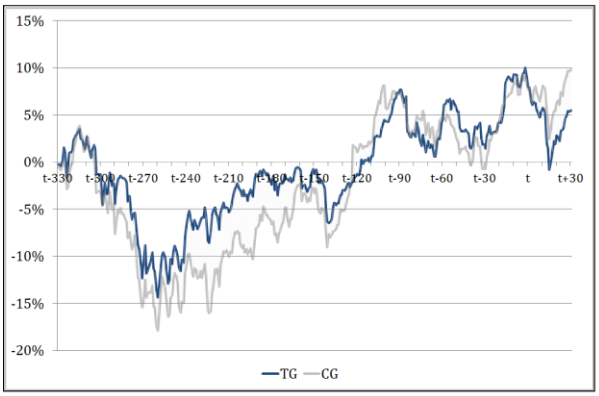


EY-II

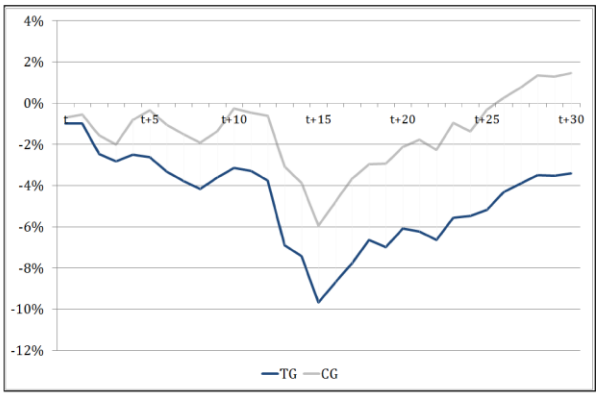


Appendix B: Cumulative Returns – Part II

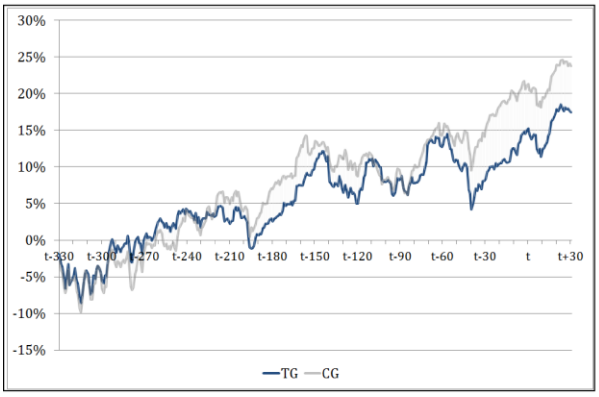
KPMG-I



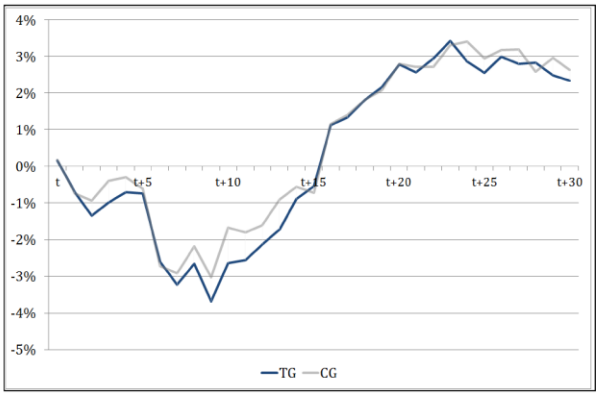
KPMG-I



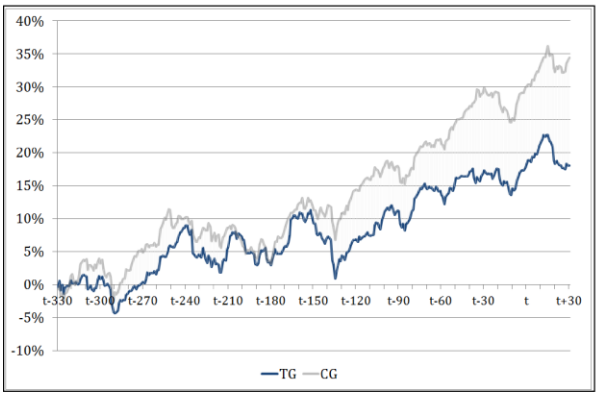
KPMG-II



KPMG-II



KPMG-III



KPMG-III