CONTENTS

FINANCING CHOICES OF INSURANCE COMPANIES: A SUMMARY OF THE LITERATURE

Jan Dhaene*, Matthias Saerens**, Frederiek Schoubben***, Cynthia Van Hulle**** and Gunther Wuyts*****

Abstract

This paper aims to provide a summary of the extant literature on the financing choices of insurers. We first take a closer look at the balance sheet of insurers and subsequently summarize general models of financing decisions and how the literature relates these to the specificities of an insurer's balance sheet. Next, traditional firm-specific factors that have been shown to be important in explaining financing choices are summed up and we go over the literate that places them within the context of an insurer. In a final part, we focus on some specific factors, that have been identified in the literature as affecting financing of insurers.

Keywords: financing of insurance companies; pecking order theory; trade-off theory

JEL codes: G22, G32

* KU Leuven, Faculty of Economics and Business, Department of Accountancy, Finance & Insurance (AFI), Naamsestraat 69, 3000 Leuven, Belgium; email: jan.dhaene@kuleuven.be; Phone: (+)32 (0)16 326750; Fax: (+)32 (0)16 326683.

** Corresponding author: K.U.Leuven, Faculty of Economics and Business, Department of Accountancy, Finance & Insurance (AFI), Naamsestraat 69, 3000 Leuven, Belgium; email: matthias.saerens@kuleuven.be; Phone: (+)32 (0)16 326726.

*** KU Leuven, Faculty of Economics and Business, Department of Financial Management (FEB@TMA), Korte Nieuwstraat 33, 2000 Antwerpen, Belgium; email: frederiek.schoubben@kuleuven.be; Phone: (+)32 (0)16 326462; Fax: (+)32 (0)16 326683.

**** KU Leuven, Faculty of Economics and Business, Department of Accountancy, Finance & Insurance (AFI), Naamsestraat 69, 3000 Leuven, Belgium; email: cynthia.vanhulle@kuleuven.be; Phone: (+)32 (0)16 326734; Fax: (+)32 (0)16 326683.

***** KU Leuven, Faculty of Economics and Business, Department of Accountancy, Finance & Insurance (AFI), Naamsestraat 69, 3000 Leuven, Belgium; email: gunther.wuyts@kuleuven.be; Phone: (+)32 (0)16 326731; Fax: (+)32 (0)16 326683.

I. INTRODUCTION

The financing choice of firms is one of the longest and most intensely debated topics in corporate finance. Since Miller and Modigliani (1958) showed that in perfect capital markets, this choice does not matter, finance scholars have attempted to explain the choices firms make in financing their operations. Overviews of this vast literature can be found in Harris and Raviv (1991) and Frank and Goyal (2008) among others. From these overviews, it can be noted that most of the attention has been put on industrial firms, financial institutions in contrast have been much less studied. In particular, banks and insurance companies are often deleted from the samples, which is typically motivated by the argument that their balance sheet is subject to severe regulation and moreover that their accounting rules differ from firms (Baranoff *et al.*, 2008). In our article, we aim to provide a summary of the recently emerged literature on the financing choices (i.e. the capital structure) of insurers (e.g. Cummins and Nini, 2002; Shim, 2010; De Haan and Kakes, 2010; Cheng and Weiss, 2012; Fier *et al.,* 2013). We start from the general literature in corporate finance but point, where relevant for models of financing choices, to differences between industrial firms and insurers. Also, our article can be useful for several stakeholders. Specifically, the determinants of financing choices are important for creating shareholder value since they directly affect the insurer's financing cost (De Weert, 2011), the pricing of its insurance products to customers (Osipov, 2012), and credit ratings (Van Gestel *et al.*, 2007). Policyholders should take great interest in their insurer's capital base as well since it is an indicator of resiliency in difficult situations. Moreover, given the new regulation on capital requirements in Europe, Solvency II, also regulators are interested in what influences the cost of insurers' capital.

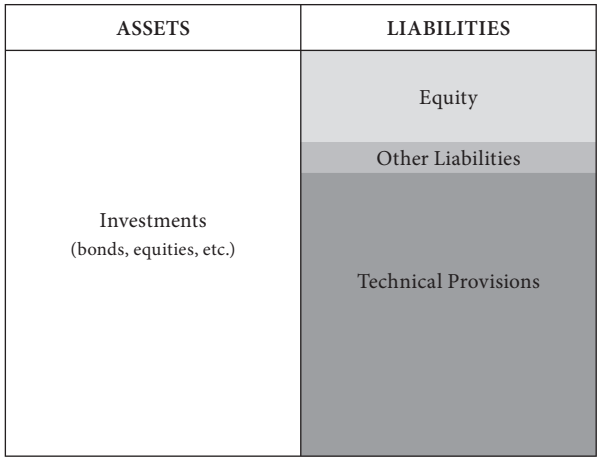
An important aspect of an insurer that is different from an industrial firm is that its business is almost completely liability driven. This stems from the fact that underwriting insurance policies and collecting premiums lead to the creation of technical provisions on the right-hand side of the balance sheet, which represent the expected amounts insurers need in order to comply with future obligations. Not surprisingly, these provisions represent the most important liability in this industry in practice. Equity capital serves as a buffer, should claims surpass technical provisions. In contrast to the focus on the conventional debt versus equity paradigm, the focus of the insurance literature is therefore on technical provisions versus equity. To sort out the implications of these differences between insurers and conventional industrial firms, we first briefly discuss the business model of an insurer. Then we summarize the two main theories as developed for financing choices of industrial firms in corporate finance (i.e. the trade-off theory and the pecking order theory), and, in a next step, use the structure of both theories to order the capital structure arguments from the insurance literature. This allows us to compare both literatures and simultaneously pinpoint to what extent the arguments of corporate finance carry over to insurance.

The remainder of this paper is organized as follows. Section II discusses an insurer's main activities and how these are translated to its balance sheet. Section III provides an overview of corporate finance theories on financing choices of firms, and we pay particular attention to literature that applies these models to the specific context of insurers. Next, section IV uses these theories to infer the firm-specific characteristics that influence financing. Section V concludes.

II. AN INSURER'S BALANCE SHEET

The core business of an insurance company is to provide policyholders the opportunity to transfer certain risks, which they are not willing to bear themselves, to the insurer. Upon the occurrence of a pre-specified event, the insurance underwriter is then obliged to financially compensate the insured. This can take a plethora of different forms, ranging from indemnification in the case of a car accident to payouts contingent on death before a certain age in the case of term life insurance. However, in exchange for underwriting these risks and providing policyholder compensation, the insurer demands an *upfront* premium. This implies that he only learns whether premiums are sufficient to cover the stochastic claim payments after the risky events have materialized. The possibility that in a certain year actual claims will outweigh expected claims, on which premium rates are based, thus constitutes a major risk for the insurer.[1] When such an adverse situation presents itself, equity comes into play as a buffer to absorb unexpected losses (Doff, 2011).

Figure 1. Simplified balance sheet of an insurance company



[1] Note that actual premiums do not only contain a part to compensate for expected claims but also contain a cost component and a profit charge (Doff, 2011).

This business model deviates noticeably from that of other (non-financial) industries. One key difference is that the roles of customer and financier are merged here, since policyholders supply financing to the company through their premium payments. Because the insurer becomes indebted to the policyholders once premiums are collected, customer interaction happens with the liability side of the balance sheet.[2] By contrast, in a typical industrial enterprise operations are asset driven and customers do not become debt holders. This is illustrated in Figure 1, where the promises granted by the insurance company appear as technical provisions (or reserves) on the right-hand side of the balance sheet.[3] These provisions can be thought of as the expected outlays for claims and hence can be regarded as expected amounts owed to the policyholders; in practice technical provisions represent the single most important balance sheet item on the liability side for an insurer. It might be argued that they are analogous to a bank's deposits, were it not that technical provisions are merely best estimates of future pay outs while deposits are liabilities of which the total repayment amount is certain (De Weert, 2011). Further, the nature of the technical provisions is dependent on the type of insurance product it is linked to. A common segmentation is made between sellers of life and non-life type products. As noted by Lencsis (1997), the former segment's technical provisions consist of 'policy reserves'. These are defined as the present value of future benefits minus the present value of future premiums, emphasizing the long-term character of this contract type. In the non-life industry, one rather uses the term 'loss reserves'- which relates mainly to incurred losses – and 'unearned premium reserves'- which is the part of the premiums for which insurance coverage has not yet been provided. Combining the concepts above, we can now identify a first source of insurance profits: the underwriting profit. This is equal to the net premiums written (gross premiums minus reinsurance costs) minus claim expenses (i.e. claim payments as well as claim settlement costs such as fees paid to experts) minus net changes in technical provisions.

In the period between premium reception and claim payout, insurers can invest the disposable funds at their discretion. The asset side typically consists of financial investments, primarily money market instruments, bonds, equities, mortgages and some real estate. The return on these investments provide the insurer with a second profit source: the investment profit. Investment decisions are made with prudence, however, since the investment portfolio ultimately serves to compensate policyholders when necessary (CEA, 2010). To further

2 In line with this reasoning, Baranoff *et al.* (2008) state that insurance policies can be looked at as contingent debt. The premiums then represent the loan proceeds which have to be repaid by the insurer, contingent on the occurrence of a loss. The debt's face value is stochastic in this case since actual payouts to a single individual tend to deviate widely from the premiums collected from that individual. Still, from a legal perspective, the policyholder does not become a creditor before an actual loss occurs.

3 This figure is a simplification of reality as we make abstraction of other potential financing mechanisms such as wholesale funding or hybrid securities and of other assets such as fixed-assets but also goodwill, deferred acquisition costs, etc.

safeguard policyholders' interests, regulatory requirements typically enforce that the asset portfolio should cover the technical provisions. In addition, liabilities (i.e. the policies) of an insurer are typically quite illiquid whereas the majority of its assets are liquid, resulting in a favorable liquidity position. Moreover, consistent with the idea of sound asset-liability management (ALM), assets are matched with the contracts found on the liabilities side. Duration matching plays a key role in the life insurance industry especially because of the long-term nature of their engagements. As a final remark, note that a substantial amount of the premiums received is not invested in the investment portfolio but rather ceded to reinsurance companies to keep overall risk levels manageable.

In practice most of the extra financial assets above technical provisions is financed by equity capital, called 'surplus' in insurance jargon. It acts as a buffer to absorb unexpected losses in the insurer's main activities, i.e. underwriting and investing.[4] A strong internal motivation for holding equity is therefore its assurance of business continuity, which makes it indispensable to reap profits in the future.[5] However, insurance firms also have a strong incentive to reduce equity as using equity entails several costs like double taxation. Specifically, as additional equity is typically invested in financial assets, insurance firms need to "beat the financial markets" in their portfolio management to create a positive net present value on their investment activities. Ample empirical evidence on this subject has shown that consistently outperforming financial markets is very difficult. Furthermore, investing activities create costs (administration, trading, …) and investment income of insurers is typically subject to corporate taxation. As a result, keeping unnecessary equity for loss absorption comes at a cost as it implies extra investments in projects which are likely to have a negative net present value (i.e. extra portfolio investments in financial assets). Therefore, in order to enhance the equity position of insurers, regulators have imposed explicit solvency rules which aim to protect policyholders and ensure a stable economic system as a whole (Doff, 2011). Prime examples are the Solvency framework in the European Union and the system of Risk Based Capital (RBC) in the United States.

The final category on the balance sheet, other liabilities, is a residual class consisting of diverse liabilities such as bank borrowings, other current liabilities, other long-term liabilities and inter-company liabilities. It is important to stress that policyholders are first in line to the remaining assets should the company default, making (most) other liabilities subordinated to the insurance liabilities.

4 Equity capital is only needed for *un*expected losses since expected losses (from underwriting activities) are supposed to be priced into the products in the first place.

5 It also makes equity indispensable to get a good evaluation by external rating agencies (Van Gestel *et al.*, 2007).

III. CAPITAL STRUCTURE THEORIES

A. SUMMARY OF CORPORATE FINANCE MODELS

The topic of capital structure has been widely debated among researchers and remains subject to considerable disagreement. After seminal work by Miller and Modigliani (1958) showed that capital structure becomes irrelevant to firm value in perfect capital markets, various theories emerged. Often, the idea was to examine the impact of introducing one or more capital market imperfections. As a consequence, real world frictions such as corporate taxation or asymmetric information have been used to argue why capital structure should matter after all.[6] An elaborate survey of this strand of literature can be found in Harris and Raviv (1991), Myers (2001) and Frank and Goyal (2008). However, two theories are advocated in particular: the trade-off theory and the pecking order theory.[7] Here we will briefly review these theories, which are developed with the typical debt-equity financed company in mind.

1. Trade-off theory

The basic idea behind this theory is that firms pursue a target financial mix that trades off the benefits and costs of leverage and alter their debt or equity positions until the marginal benefits and costs are equal. Following their paper on capital structure irrelevance, Miller and Modigliani (1963) show that companies should maintain 100 percent debt financing when corporate tax regulation allows for the deductibility of interest charges. Hence, shielding income from taxes constitutes a significant advantage of debt, making capital structure no longer irrelevant (see Graham, 2003 for an overview of tax based theories). To mitigate the extreme result of pure debt financing, though, the literature has come up with offsetting costs. Classical expositions by Kraus and Litzenberger (1973) and Bradley *et al.* (1984), for example, introduce models where the tax advantages of leverage are balanced with the expected costs of bankruptcy.

Jensen and Meckling (1976) extend the framework of costs and benefits with the 'agency' perspective. More specifically, they add to the tax benefit vs bankruptcy trade-off above. Their theory's premise is that manager-shareholder and shareholder-debt holder conflicts are inherent to any company. The conflict of interest between managers and shareholders is essentially a moral hazard problem stemming from the fact that managers hold less than 100 percent of the residual claim (i.e. of equity). They bear the whole cost of foregoing

[6] Nevertheless, some assumptions can be relaxed without changing the irrelevance proposition, e.g. Stiglitz (1969, 1974) or Baron (1974).

[7] This implies that some theories are only discussed briefly (e.g. signaling motives as in Ross (1977)). Others are just not applicable, Cheng and Weiss (2012) argue for instance that the market timing theory by Baker and Wurgler (2002) is not appropriate for insurers because many firms in this industry are not publicly traded.

perquisites such as luxury cars or fancy offices, but only collect a part of the associated gains. If the absolute investment of the manager in the company is held constant, substituting equity for debt would increase the manager's shareholdership fraction and therefore align its interests with the external shareholders (Harris and Raviv, 1991). Furthermore, the free cash flow hypothesis by Jensen (1986) also states that taking on leverage can mitigate agency conflicts: because the money needed to make repayments of the debt reduces equity cash available to managers and thus their opportunity to overinvest. In sum, increasing leverage hampers the desire and opportunity to engage in negative net present value projects. Nevertheless, too much debt is not optimal either. It aggravates the shareholder-debt holder conflict, which arises because debt holders' claims to assets have legal priority over the shareholders' claims. This implies that shareholders only have a residual claim and only receive the market value of the firm that remains after all debt holders are paid off. This unique relationship between shareholders and debt holders causes two agency problems, i.e. asset substitution and debt overhang (see e.g. Monda, 2013). These will be discussed in more detail later in the specific context of insurance companies.

2. Pecking order theory

Influenced by Donaldson's (1961) observations on financing preferences, Myers (1984) and Myers and Majluf (1984) are the seminal contributions putting forward the pecking order theory. The latter posits that firms prefer internal over external financing and, when all internal financing is depleted, debt over equity. It is based on the assumption that managers know more about the firm's value of assets in place and about its investment opportunities than the investors, i.e. asymmetric information is present. This causes new equity issues to be underpriced so severely (reflecting average project quality) that positive net present value projects may not get financed at all (Danthine and Donaldson, 2005). The underlying logic is that managers, acting in the best interest of the current shareholders, refrain from issuing equity because the new investors may capture more than the net present value of the project, thus diluting shareholders in place. Myers and Majluf (1984) argue that this underinvestment problem can be mitigated by using securities that are less susceptible to underpricing. As a result, a 'pecking order' arises where sources of financing are prioritized according to their relative costs. Retained earnings are used up first, followed by 'safe' debt (i.e. debt where problems of asymmetric information remain limited). Once these sources are exhausted, companies move on to riskier debt types that display more equity-like features. Finally, equity is used as a last resort when debt capacity is reached. This implies that capital structure is determined by the accumulation of past financial requirements and not by an attempt to reach a target debt-equity ratio as in the trade-off theory (Myers, 2001; Shyam-Sunder and Myers, 1999).

B. APPLICABILITY TO THE INSURANCE INDUSTRY

In the context of insurers, technical provisions are added to the preceding capital structure framework of debt and equity. Hence, since insurers and traditional firms differ substantially, the trade-off and pecking order theory need to be translated to the insurance setting. To that end, one has to keep in mind that attracting funds to finance particular assets is not the ultimate goal of an insurer. Underwriting policies and creating the technical provisions that go with it, however, is. The asset portfolio is therefore a by-product, albeit an important one, of granting insurance. Overall, it is not a matter of finding the appropriate capital structure to finance certain asset-related operations, but to hold the appropriate proportion of equity (serving as a safety net) to technical provisions (representing the core business).[8] This is also reflected in the observation that the amount of other liabilities is marginal in comparison to technical provisions and only accounts for a small percentage of the balance sheet total. In fact, in the theoretical and empirical capital structure literature on insurance, technical provisions take over de role of debt within classical industrial firms since the three main parties in the insurance industry are shareholders, managers and policyholders while other debt holders only play a limited role (Pottier and Sommer, 1997).

We will therefore define leverage for an insurer as the ratio of technical provisions plus other liabilities (again, the latter are small compared to the former) divided by equity, instead of the debt-to-equity ratio commonly used in corporate finance research. Another popular leverage measure used in the insurance industry is the ratio of net premiums written to equity (Cummins and Nini, 2002). However, we will focus on the former measure for comparability reasons with the previous subsection and because prior literature claims that the latter measure is flawed (see Fier *et al.*, 2013 for details).

1. Trade-off theory

Presently we organize our summary of the arguments from the insurance literature according to the framework of the trade-off theory, for ease of comparability and in order to evaluate how they fit into the latter theory.

We turn to the trade-off between bankruptcy and tax first. As discussed above, more equity reduces bankruptcy risk and the associated costs and enhances the chances of continuing profitable operations. Conversely, a deteriorated (relative) equity position increases the risk of financial distress, which is even more important for insurers than it is for traditional companies. This is because, among other implications, default risk and insurance prices are inversely related (Cummins and Danzon, 1997; Sommer, 1996) and breaches of certain solvency thresholds trigger rating downgrades and regulatory intervention. However, holding equity can have negative consequences as well since it is usually invested in a portfolio

[8] Bear in mind that equity also serves to absorb losses in the investment portfolio.

of financial securities which entails management costs and corporate taxation of the investment income. Furthermore, declaring more policy reserves is advantageous since claim expenses and net changes in reserves can immediately be subtracted from taxable income (Bradford and Logue, 1999).[9] Hence, at heart a similar tax versus bankruptcy argument as in traditional firms continues to hold.

Next, we turn to agency considerations. While agency costs of equity, arising from conflicts between managers and shareholders, are equally applicable to insurers, the reasoning underlying the agency costs of debt alters. This is because the role of financial debt holders is much less important than it is for an industrial company, so that costs of 'debt' now concentrate on the interplay between shareholders and policyholders. Two important cases of shareholder-policyholder problems are debt overhang and asset substitution.

In times of financial difficulties, it might benefit all stakeholders to attract new equity to stabilize the insurance company. However, according to the debt overhang or underinvestment problem identified by Myers (1977), they may refrain from doing so. Reason being that funds of the new shareholders will serve mainly as a bail-out for policyholders in place (De Weert, 2011). Shareholders are thus unlikely to inject capital when it's needed the most. This issue can be alleviated by scaling down leverage or renegotiating contracts when possible. Culp (2011) suggests the use of reinsurance as an alternative, which he claims is comparable to a 'synthetic' equity infusion. Another solution could be to use hybrid securities, e.g. contingent convertibles, which are bonds that transform into common equity when a certain trigger is hit (such as a breach of solvency requirements).

Next, we turn to the issue of asset substitution (also called risk-shifting) as worked out in the seminal contributions of Jensen and Meckling (1976) and Galai and Masulis (1976). This theory starts from the observation that shareholders have limit liability and also are the residual claimants of the assets of the firm (i.e. they obtain the remainder of the assets after all policy holders are paid). Consequently, the equity can be seen as a European call option with strike price equal to the value of the liabilities (see also Merton, 1974). Hence, managers who act in the interest of the shareholders have an incentive to select projects of higher risk, especially when leverage is high and firm value is low (Macminn, 1987). The reasoning is that the value of a call option is positively related to volatility. Such projects typically return a very high gain when successful but only have a low success rate. When the project is successful, shareholders benefit from it. However, if the project goes bad, the policyholders have to bear most of the costs because of the shareholders' limited liability. Again, deleveraging is a way to mitigate this conflict.

[9] Gatzert and Schmeiser (2008) further examine the interactions between taxes and capital structure. They claim that insurers alter prices as well as capital structures in response to changes in tax regime. Their model predicts that premiums are increased by the present value of tax payments to keep safety levels equal, which leads to greater leverage.

2. *Pecking order theory*

As argued by Cummins and Nini (2002), asymmetric information between investors and managers of insurers plays a key role for the implications of the pecking order theory. Commercial firms typically only face asymmetric information in their assets, but insurance companies face it for both assets and liabilities (Zhang *et al.*, 2009). In fact, asymmetric information seems even larger for insurers than for banks, when looking at the disagreements in rating agencies (Morgan, 2002). Nevertheless, since the assets of an insurer typically are composed of marketable securities, the asymmetric information problem on the asset side seems relatively small. In contrast, the technical provisions are much more difficult to assess for investors, inducing a much more opaque liability side of the balance sheet. The reasoning is that insurers have a significant amount of flexibility in computing these provisions (Cummins and Nini, 2002). Moreover, some regulatory frameworks do not require to disclose granular information about the covered risks (e.g. Solvency I).[10] The liabilities therefore represent the most important source of asymmetric information in this setting. As a result, in practice, if insurers seek extra financing, they usually focus on internal resources as this funding signals the least information to investors.[11] Next follow other liabilities and lastly new equity issues.

Further, Myers and Majluf's (1984) ideas are also used to shed light on the underwriting cycle, i.e. the cyclical manner in which profitability in the property-casualty (i.e. non-life) industry tends to rise and fall. A popular explanation is the capacity constraint model by Winter (1988) and Gron (1994). This model assumes that capital does not flow freely into and out of the insurance industry, losses are correlated across policies (leading to industry-wide shocks) and insurance supply depends on the insurer's capital base (Weiss and Chung, 2004). According to the model, insurers will react to a large capital shock (e.g. due to unexpected losses) by rebuilding capital internally instead of issuing costly equity. Since it is assumed that insurers need sufficient equity to keep the probability of bankruptcy low, to maintain the ability to meet policyholder claims and to comply to regulatory requirements, the diminished equity will lead to a reduction in capacity. This is often combined with heightened demand for insurance following the shock, ultimately resulting in soaring insurance prices and profitability (Doherty *et al.*, 2003). Competition will then enter the market and an increased supply is likely to suppress premium rates, paving the way for a subsequent shock and starting

[10] Note that the difficulty in judgment, combined with the fact that changes in technical provisions have implications for the income statement, implies that insurers have considerable earnings management opportunities.

[11] This argument is related to the signaling theory, which is based on information asymmetries between investors and managers (as in the pecking order theory). According to Ross (1977), managers can create a signal that influences the market's perception about future cash flows by making changes to the capital structure. As in the pecking order theory, using internal resources as a funding mechanism signals little information since it is surrounded by few adverse selection issues.

off the cycle once again. Weiss (2007) states that, consequently, insurers hoard capital in times of excess capacity (when capital is readily available) in anticipation of tight markets, which is in line with the pecking order theory. Winter (1994), Gron (1994) and Weiss and Chung (2004) find empirical evidence consistent with the capacity constraint theory.

Most empirical studies on the capital structure of insurers provide evidence that insurers maintain a target capital structure, although the evidence also points towards the presence of pecking order elements (Harrington and Niehaus, 2002; De Haan and Kakes, 2010; Shim, 2010; Fier *et al.*, 2013; Cheng and Weiss, 2012). In fact, such findings are very much in line with the empirical results for traditional firms (e.g. Titman and Wessels, 1988; Rajan and Zingales, 1995 and Frank and Goyal, 2009). Some more detail is offered in Section IV below, where we discuss several important firm-specific determinants of capital structure.

IV. FIRM-SPECIFIC DETERMINANTS OF CAPITAL STRUCTURE

Building on the preceding theories, several authors have examined firm-specific factors that influence a traditional company's capital structure (e.g. Titman and Wessels, 1988; Rajan and Zingales, 1995 and Frank and Goyal, 2009). Among the more commonly cited ones are firm size, profit, growth opportunities, risk and lagged capital structure. In the next paragraph, an overview is given of the applicability of these traditional determinants to the insurance industry. We conclude with some insurance-specific factors that have been thoroughly discussed in the literature: product market interactions, organizational form, firm diversification and institutional setting and regulatory requirements.

A. TRADITIONAL DETERMINANTS

Firm size. Because a large number of insured risks in the pool makes losses more predictable, a large insurer needs less capital to achieve its desired insolvency risk (Cummins and Nini, 2002). Their increased access to capital markets also requires them to hold less ex-ante capital. Nevertheless, as large insurers are generally surrounded by less asymmetric information (since they tend to be more mature, receive more attention from financial markets and regulators, etc.), the pecking order theory dictates that they will be less levered because raising external capital is cheaper for them. Empirical research mainly supports the first line of thought (e.g. Fier *et al.*, 2013; De Haan and Kakes, 2010).[12]

Profitability. Consistent with the pecking order theory, profitable insurers have more internal funds available which they can hoard as a buffer for future use (Harrington and

[12] Fier *et al.* (2013) examine a panel data sample consisting of U.S. affiliate non-life insurers for the period 1996–2009. De Haan and Kakes (2010) study a panel data sample of Dutch life and non-life insurers during 1995–2005.

Niehaus, 2002). In this way, equity issues can be avoided when adverse shocks (e.g. catastrophes or a plummeting stock market) deplete capital. This view is supported by the capital constraint theory explained above and suggests a positive relation between profitability and capital (Cheng and Weiss, 2012). In contrast, the trade-off theory hypothesizes a negative relation since increasing capital aggravates the problems associated with free cash flow (Jensen, 1986). Results of empirical studies are mostly in line with the predictions of the pecking order theory (e.g. Shim, 2010; Harrington and Niehaus, 2002).[13]

Growth opportunities. The effect of this firm-specific factor is ambiguous according to prior findings (e.g. Fier *et al.,* 2013; Cummins and Nini, 2002).[14] On the one hand, the trade-off theory suggests a positive relation with capital as, at high debt levels, growth exacerbates agency conflicts of debt related to underinvestment (e.g. reduces flexibility in accessing future investment opportunities) and mitigates agency conflicts of equity (reducing free cash flow issues) (Titman and Wessels, 1988; Frank and Goyal, 2008; Cheng and Weis, 2012; Shim, 2010). The pecking order theory predicts a similar outcome: firms with more growth opportunities will have higher equity levels to avoid raising costly capital (Cummins and Nini, 2002). On the other hand, past growth may have enhanced debt (i.e. technical provisions) and thereby enhanced leverage.

Risk. The literature, and especially the one on financial firms, considers risk to be an important capital structure determinant. In the insurance industry, a distinction should be made between asset risk and product risk, referring to investment and underwriting activities respectively. The former is related to holding a portfolio of risky assets while the latter stems from engaging in risky lines of business. To unveil how these risks relate to capital, we first turn to the option pricing methodology for insurance developed by Cummins (1988) and Cummins and Sommer (1996). Consider a simple one-period, two-date model where policies are issued at t = 0 and claims are paid out at t = 1. The idea is that policyholders' claims are fully settled on t = 1 when the value of the assets (V) is larger than the value of the outstanding liabilities (L). However, they do not receive more than V in case the asset value has dropped below the value of total liabilities. This implies that the end-of-period payoffs in both states are L and V respectively, i.e. the liabilities pay off min(L,V) which is equivalent to L-max(L-V,0).[15] On t = 0, the market value of insurers' promises to policyholders therefore equals the present value of liabilities (i.e. similar to the present value of a riskless bond) minus an insolvency put option on the company assets (Sommer, 1996; Phillips *et al.*, 1998). Shim

[13] Shim (2010) investigates a panel data sample that includes U.S. non-life insurers during 1993–2004. Harrington and Niehaus (2002) use a panel data sample consisting of U.S. non-life insurers for the period 1991–1998.

[14] Cummins and Nini (2002) study a panel data sample that contains U.S. non-life insurers for the period 1993–1998.

[15] Note that this reasoning assumes a corporate form with limited liability on behalf of the shareholders. Also notice that when a guarantee fund is present, the liabilities are worth L to the policyholders in all states since the guarantee fund assumes the insolvency put option itself (Merton and Perold, 1993).

(2010) explains that the value of the option captures the insolvency risk of the company and is a function of firm value, loss liabilities and risks. Hence, default risk is jointly determined by risk in the insurer's asset-liability portfolio and by capital structure levels.

The preceding discussion implies that risk (asset or product related) should be offset by increased capitalization to keep the probability of default at acceptable levels.[16] Results by Guo and Winter (1997) indicate that capital is indeed positively related to the degree of uncertainty in insurance losses (i.e. to product risk).[17] Further, De Haan and Kakes (2010) make clear that actual solvency margins are a function of risk even when capital requirements are not risk-based, as is the case for Europe's Solvency I. However, capital structure can influence risk as much as risk can influence capital structure. It seems plausible that a well-capitalized insurer has more degrees of freedom to increase risks in the investment portfolio while still maintaining an appropriate probability of default. It can thus be hypothesized that equity and risk adjustments occur in the same direction. Shim (2010) finds confirmation for this using data on U.S. property-liability insurers and properly accounts for potential endogeneity problems (due to simultaneous causality between risk and equity) by adopting an instrumental variables approach. Yet, Baranoff and Sager (2002) point out that it is necessary to distinguish between asset and product risk since they have a different impact on equity. Note that the probability of default can be kept at reasonable levels via alternative ways as well, though. Shiu (2011) for example postulates that firms who purchase more reinsurance can hold less equity since reinsurance reduces the strain on capital. When the cost of reinsurance (i.e. the reinsurance premium) is cheaper than external financing, the insurer is therefore expected to rely more on reinsurance.

Lagged capital structure. As already mentioned above, several authors have found evidence supporting the notion that insurance firms have a target capital structure. In practice, however, actual capital structure levels often deviate from their target. One expects, just as in the case of a traditional firm, that insurers do not revert back to their optimal level immediately but rather opt for gradual reversion, since capital structure adjustments are costly (Myers and Majluf, 1984). Combined insights from the trade-off and pecking order theories – target capital and costly adjustment, respectively – thus allow to explain the insurer's behavior in a multi-period setting. These ideas are bundled under the 'dynamic' trade-off theory, which postulates that the previous period's capital structure impacts the one of today. Shim (2010) finds that US property-liability insurers do have a target capital structure, but revert to it very slowly. He estimates a speed of adjustment of around four percent, which he attributes to high adjustment costs and illiquid markets. Fier *et al.* (2013) also conclude that insurers have target capital structures and report that internal capital markets of a business group are one of the major ways to reduce deviations from it. Specifically,

16 This is also implied by the trade-off theory.

17 Guo and Winter (1997) also test their theoretical model on a panel data sample of U.S. non-life insurers over the period 1990–1995.

their partial adjustment model indicates that affiliate reinsurance transactions are being used to change the level of written premiums and the potential liability associated with ceded risks.

B. INSURANCE SPECIFIC VARIABLES[18]

1. Product market interactions

Life and non-life insurance. While non-life insurance offers compensation when a previously agreed upon incident occurs, life insurance promises future payments on death or at a certain age (Doff, 2011). The coverage period of life insurance is therefore much longer than of non-life insurance and, moreover, life-insurance is generally more predictable (Hull, 2012). This implies that capitalization in the life industry should be substantially lower than in the non-life industry, which is observed in practice as well.

Personal and commercial lines of business. Policyholders are sensitive to the insolvency risk of their insurance company, often caused by low capitalization levels. This implies that the demand for insurance decreases once capital buffers become insufficient, especially when the policyholders are insufficiently protected by the asset portfolio or by a guarantee fund (De Haan and Kakes, 2010). Cummins and Danzon (1997) state that safer insurers can thus charge higher prices. Further, the demand for insurance can be subdivided into demand for commercial insurance and demand for personal insurance. Commercial buyers generally have lower switching costs, base themselves more on financial ratings and have better knowledge about the insurer's financial status. Cummins and Nini (2002) argue that, as a consequence, leverage will be lower if the fraction of corporate insurance lines relative to personal lines increases.

Long-tail lines of business. When a firm engages in long-tail lines of business, i.e. lines where the time lag between premium payment and claim settlement is substantial, the funds are under management's supervision for an extended period of time. This offers managers the possibility to exhibit opportunistic behavior by engaging in activities that give them private benefits but are not in the best interest of policyholders (Cumming and Nini, 2002). Diminished equity levels and the ensuing increased performance pressure can then provide the correct incentives by discouraging managers from acting in such a way.

2. Organizational form

Insurance companies are often classified according to their organizational form: mutual or stock insurer. Whereas shareholders and policyholders are distinct stakeholders in the stock type, their roles are unified in a mutual company. Hence, the organizational form directly

[18] ???

impacts the way an insurer is subject to agency costs (e.g. Mayers and Smith, 1981; Fields and Tirtiroglu, 1991; Pottier and Sommer, 1997). In fact, conflicts between shareholders and policyholders are resolved by definition within a mutual, which suggests that these organizations need less equity to mitigate such problems (Lamm-Tennant and Starks, 1993). Additionally, less equity is also justified because shareholder-manager conflicts are likely to be worse for a mutual insurer as fewer mechanisms are available for shareholders to exert control over managers (Cummins and Nini, 2002). Conversely, Harrington and Niehaus (2002) and Froot and Stein (1998) state that mutuals may hoard capital for precautionary reasons, due to their limited access to capital markets in comparison to stock insurers.[19] The predicted relation between equity and having the mutual shareholdership form is therefore ambiguous.

3. Firm diversification

Insurance companies can mitigate their risks by diversifying operations, allowing them to maintain a lower solvency ratio (Cummins and Nini, 2002; Klein *et al.*, 2002).[20] Three channels through which diversification can be achieved are subsequently discussed. First, insurers can engage in product line diversification. By combining imperfectly correlated insurance products such as life and non-life insurance, personal and commercial lines of business and long and short tailed lines of business, risks (and hence capital) can be lowered. Shim (2010) uses a product line Herfindahl index as a measure for product diversification and shows that it indeed decreases the need for capital. Cummins *et al.* (2010), however, study product diversification in the U.S. life and non-life industry during 1993–2006 and conclude that insurers who strategically focus on a small set of core products outperform insurers who offer a wide product range. Second, diversification can occur across geographical locations. This leads to lower correlations across claims (especially for contracts that insure risks bounded by a geographical location, such as earthquakes), which makes it possible to reduce capital. A third plausible avenue for diversification stems from combining insurance activities with other financial and non-financial operations (i.e. to create a financial conglomerate). A relevant type of financial conglomerate in the context of this paper is the bancassurance model, in which insurers can sell their products via a partnering bank's clientele. Kuritzkes *et al.* (2003) argue that such strategic diversification choices greatly affect the level of capital and develop a novel approach to aggregate risks in a bank-insurance conglomerate. Finally, note that diversification can also affect systemic risk. This can have material capital structure implications since it might influence the ability to attract new capital and new customers. Slijkerman *et al.* (2013) use extreme value analysis on a sample of stock returns for ten large

[19] De Haan and Kakes (2012) note that raising equity is not straightforward for a mutual because its most important source of capital, i.e. the shareholders, are also the policyholders.

[20] Klein *et al.* (2002) investigate a cross-sectional data sample including all U.S. non-life insurers during 1997.

European banks and insurers during 1992–2003 and provide evidence that insurers exhibit higher interdependencies than banks and that the interdependency across the two sectors is low, suggesting that financial conglomeration can reduce systemic risk.[21]

4. Institutional setting and regulatory requirements

A final insurance specific variable discusses the importance of the institutional setting and regulatory requirements (see e.g. Laeven and Periotti, 2010). Historically, insurance firms all over the world have been regulated strictly with the objective to avoid excessive insurance premiums. Over time, however, supervisors have deregulated premium rates and focused attention towards regulating technical provisions and minimum capital requirements (Klein *et al.*, 2002).[22]

The current regulatory framework for European insurers, i.e. Solvency I, imposes minimum capital requirements which depend on the technical provisions (life insurance) or on the maximum of earned premiums and gross claim experience (non-life insurance) (see De Weert, 2011). Although Solvency I is only in place since 2002, its structure is based on the initial Insurance Directives which date back to the 1970s and which are not risk-sensitive. As a consequence, risk reductions do not always lead to reductions in the requirements (e.g. increasing non-life premiums increases capital requirements instead of decreasing them) and certain risks are not addressed explicitly (e.g. market risks) (Doff, 2011). The upcoming and updated framework Solvency II tries to overcome these shortcomings by adopting a risk-based approach, analogously to the three-pillar framework for banks Basel II. Nonetheless, there are still few studies available that document the possible effects of a transition towards such a risk-based system. De Haan and Kakes (2010) offer preliminary evidence that the adjustment to Solvency II will be relatively smooth because, for the 350 Dutch insurers in their sample, current non-risk based capital requirements seem to be non-binding (i.e. the majority holds more capital than legally required). Note that other countries have introduced different solvency systems, with the majority applying at least some basic form of risk weighting. Important examples include risk based capital (RBC) in the U.S. and the Swiss Solvency Test (SST) in Switzerland.

Aside from the direct effect regulation has on capital structure, it can also have a more indirect impact. As a case in point, consider the consequences of imposing rules on agency conflicts. When the obligations to publicly disclose information are tightened, for example, the degree of asymmetric information between policyholders and managers is expected to

[21] Somewhat in contrast with these findings, however, are the results of Cummins and Weiss (2014). They conclude that the core activities of U.S. insurance companies do not pose systemic risk. Nonetheless, when insurers engage in noncore activities (such as banking activities), systemic risk might become a concern.

[22] Note that such deregulation generally has a non-negligible effect on a firm's operating environment and its leverage decisions (Ovtchinnikov, 2010).

change. Since agency conflicts strongly influence the capital and risk decisions of the insurer (Cummins *et al.*, 2010), this may affect the capital structure of the insurance company.

Clearly, supervisors have many (direct and indirect) channels at their disposal to influence capital structure decisions and, by extension, the overall soundness of the insurance company. The findings of Pasiouras and Gaganis (2013) provide guidance with respect to the most efficient regulatory channels. More specifically, they find that regulation related to technical provisions and investments impacts the soundness of the insurer whereas corporate governance, internal control rules and capital requirements have no impact.[23]

V. CONCLUSIONS

Although capital structure is a topic that has been examined extensively in the corporate finance literature, financial firms are almost always excluded from the discussion. This article aims to complement the traditional corporate finance literature by providing a summary of the recent contributions that have focused on financing choices of insurance firms.

As opposed to industrial firms, the insurance companies' business is almost entirely liability driven. This stems from the fact that technical provisions have to be build up out of policyholders' premium payments, in order to comply with future obligations. Not surprisingly, provisions represent the most important liability in this industry and equity capital serves as a buffer, should these provisions not suffice to pay out claims. In contrast to the conventional debt-equity paradigm, the focus in the literature has therefore been on technical provisions and equity. As a consequence, the trade-off theory – which argues that firms have a target capital structure that balances the benefits and costs of leverage – is now about finding the optimal mix between equity and technical provisions. Furthermore, while the idea behind the pecking order theory remains valid (i.e. asymmetric information between managers and investors causes new equity issues to be expensive, creating a hierarchy in financing instruments), asymmetric information is mainly about the appropriateness of technical provisions and not about asset-related uncertainty anymore. Finally, next to insurance-specific factors such as organizational form (i.e. being a mutual or a stock insurer), product market interactions (i.e. life versus non-life insurer, personal versus commercial lines of business and long versus short-tail lines of business), firm diversification and regulatory requirements, empirical evidence indicates that – likewise industrial firms – the influence of determinants as firm size, profit, growth opportunities, risk and lagged capital structure continues to be substantial.

[23] The study by Pasiouras and Gaganis (2013) compromises a panel data sample including life and non-life insurers from 46 countries over the period 2005–2007.

REFERENCES

Baker, M., and Wurgler, J. (2002), "Market Timing and Capital Structure", *Journal of Finance,* 57(1): 1–32

Baranoff, E., Sager, T.W., and Shively, T. (2008), "Rebalancing Target Capital in the Financial Sector: The Case of Life Insurance", McCombs Research Paper Series No. IROM-02–09.

Baranoff, E.G., and Sager, T.W. (2002), "The Relations Among Asset Risk, Product Risk, and Capital in the Life Insurance Industry", *Journal of banking & finance*, 26(6): 1181–1197.

Baron, D.P. (1974), "Default Risk, Homemade Leverage, and the Modigliani-Miller Theorem", *American Economic Review*, 64(1): 176–182.

Bradford, D.F., and Logue, K. (1999), "The Influence of Income Tax Rules on Insurance Reserves", K. FROOT (ed.), *The Financing of Catastrophe Risk*, University of Chicago Press.

Bradley, M., Jarrell, G.A., and Kim, E. (1984), "On the Existence of an Optimal Capital Structure: Theory and Evidence", *Journal of Finance*, 39(3): 857–878.

Browne, M.J., Carson, J.M., and Hoyt, R.E. (1999), "Economic and Market Predictors of Insolvencies in the Life-Health Insurance Industry", *Journal of Risk and Insurance,* 66(4): 643–659.

CEA Insurers of Europe (2010), "Insurance: A Unique Sector. Why Insurers Differ from Banks", Brussels, Belgium.

Cheng, J., and Weiss, M.A. (2012), "Capital Structure in the Property-Liability Insurance Industry: Tests of the Tradeoff and Pecking Order Theory", *Journal of Insurance Issues,* 35(1):1–43.

Culp, C.L. (2011), "Structured Finance and Insurance: The ART of Managing Capital and Risk", Wiley.

Cummins, J.D. (1988), "Risk-Based Premiums for Insurance Guaranty Funds", *Journal of Finance*, 43(4): 823–839.

Cummins, J.D., and Danzon, P.M. (1997), "Price, Financial Quality, and Capital Flows in Insurance Markets", *Journal of Financial Intermediation*, 6(1): 3–38.

Cummins, J.D., and Nini, G.P. (2002), "Optimal Capital Utilization By Financial Firms: Evidence From the Property-Liability Insurance Industry", *Journal of Financial Services Research*, 21(1/2): 15–53.

Cummins, J.D., and Sommer, D.W. (1996), "Capital and Risk in Property-Liability Insurance Markets", *Journal of Banking and Finance*, 20(6): 1069–1092.

Cummins, J.D., and Weiss, M.A. (2014), "Systemic Risk and the US Insurance Sector", *Journal of Risk and Insurance (Accepted)*

Cummins, J.D., Weiss, M.A., Xie, X., and Zi, H. (2010), "Economies of Scope in Financial Services: A DEA Efficiency Analysis of the US Insurance Industry", *Journal of Banking and Finance*, 34(7): 1525–1539.

Danthine, J.P., and Donaldson, J.B. (2005), "Intermediate Financial Theory", Academic Press.

De Haan, L., and Kakes, J. (2010), "Are Non-Risk Based Capital Requirements for Insurance Companies Binding?", *Journal of Banking and Finance*, 34(7): 1618–1627.

De Weert, F. (2011), "Bank and insurance capital management", Wiley Finance, West Sussex, UK.

Doff, R. (2011), "Risk Management for Insurers", Risk Books, London, UK.

Doherty, N.A., Lamm-Tennant, J., and Starks, L.T. (2003), "Insuring September 11th: Market Recovery and Transparency", *Journal of Risk and Uncertainty*, *26*(2–3): 179–199.

Donaldson, G. (1961), "Corporate Debt Capacity: A Study of Corporate Debt Policy and the Determination of Corporate Debt Capacity", Boston, Division of Research, Harvard Graduate School of Business Administration.

Fields, J.A., and Tirtiroglu, D. (1991), "Agency-Theory Implications for the Insurance Industry: A Review of the Theoretical and Empirical Research", *Quarterly Journal of Business and Economics*, 30(1): 40–61.

Fier, S.G., McCullough, K.A., and Carson, J.M. (2013), "Internal Capital Markets and the Partial Adjustment of Leverage", *Journal of Banking and Finance,* 37(3): 1029–1039.

Frank, M. Z., and Goyal, V.K. (2009), "Capital Structure Decisions: Which Factors Are Reliably Important?", *Financial Management*, 38(1), 1–37.

Frank, M., and Goyal, V. (2008), "Trade-off and Pecking Order Theories of Debt", E. ECKBO (ed.), Handbook of Corporate Finance: Empirical Corporate Finance, Vol.2, Elsevier, Amsterdam.

Froot, K.A., and Stein, J.C. (1998), "Risk Management, Capital Budgeting, and Capital Structure Policy for Financial Institutions: An Integrated Approach", *Journal of Financial Economics*, *47*(1): 55–82.

Galai, D., and Masulis, R W. (1976), "The Option Pricing Model and The Risk Factor of Stock", *Journal of Financial Economics,* 3(1/2): 53–81.

Gatzert, N., and Schmeiser, H. (2008), "Combining Fair Pricing and Capital Requirements for Non-Life Insurance Companies", *Journal of Banking & Finance*, *32*(12): 2589–2596.

Graham, J.R. (2003), "Taxes and Corporate Finance: A Review", *Review of Financial Studies,*16(4): 1075–1129.

Gron, A. (1994), "Capacity Constraints and Cycles in Property-Casualty Insurance Markets", *RAND Journal of Economics,* 25(1): 110–127.

Guo, D., and Winter, R.A. (1997), "The Capital Structure of Insurers: Theory and Evidence", Working Paper, Sauder School of Business.

Harrington, S.E., and Niehaus, G. (2002), "Capital Structure Decisions in the Insurance Industry: Stocks Versus Mutuals", *Journal of Financial Services Research*, 21(1): 145–163.

Harris, M., and Raviv, A. (1991), "The Theory of Capital Structure", *Journal of Finance*, 46(1): 297–355.

Hull, J. (2012), "Risk Management and Financial Institutions", Wiley.
Jensen, C.J. (1986), "Agency Costs of Free Cash Flow, Corporate Finance, and Takeovers", *American Economic Review*, 76(2): 323–329.
Jensen, C.J., and Meckling, W.H. (1976), "Theory of the Firm: Managerial Behavior, Agency Costs and Shareholdership Structure", *Journal of Financial Economics*, 3(4): 305–360.
Klein, R.W., Phillips, R.D., and Shiu, W. (2002), "The Capital Structure of Firms Subject to Price Regulation: Evidence from the Insurance Industry", *Journal of Financial Services Research*, 21(1–2): 79–100.
Kuritzkes, A., Schuermann, T., and Weiner, S.M. (2003), "Risk Measurement, Risk Management, and Capital Adequacy in Financial Conglomerates", *Brookings-Wharton Papers on Financial Services*, *2003*(1): 141–193.
Kraus, A., and Litzenberger, R.H. (1973), "A State-Preference Model of Optimal Financial Leverage", *Journal of Finance*, 28(4): 911–922.
Laeven, R.J., and Perotti, E.C. (2010), "Optimal Capital Structure for Insurance Companies", Working paper, Tilburg University and University of Amsterdam.
Lamm-Tennant, J., and Starks, L.T. (1993), "Stock Versus Mutual Shareholdership Structures: The Risk Implications", *Journal of Business*, 66(1): 29–46.
Lencsis, P.M. (1997), "Insurance Regulation in the United States: An Overview for Business and Government", Quorum Books, Westport, CT.
Macminn, R.D. (1987), "Insurance and Corporate Risk Management", *Journal of Risk and Insurance*, 54(4): 658–677.
Mayers, D., and Smith JR, C.W. (1981), "Contractual Provisions, Organizational Structure, and Conflict Control in Insurance Markets", *Journal of Business*, 54(3): 407–434.
Merton, R.C. (1974), "On the Pricing of Corporate Debt: The Risk Structure of Interest Rates", *Journal of Finance*, 29(2): 449–470.
Merton, R.C., and Perold, A. (1993), "Theory of Risk Capital in Financial Firms", *Journal of Applied Corporate Finance*, *6*(3): 16–32.
Modigliani, F., and Miller, M.H. (1958), "The Cost of Capital, Corporation Finance and the Theory of Investment", *American Economic Review*, 48(3): 261–297.
Modigliani, F., and Miller, M.H. (1963), "Corporate Income Taxes and the Cost of Capital: A Correction", *American Economic Review*, 53(3): 433–443.
Monda, B., Giorgino, M., and Modolin, I. (2013), "Rationales for Corporate Risk Management: A Critical Literature Review", Available at SSRN 2203546.
Morgan, D.P. (2002), "Rating Banks: Risk and Uncertainty in an Opaque Industry", *American Economic Review*, 92(4): 874–888.
Myers, S.C. (1977), "Determinants of Corporate Borrowing", *Journal of Financial Economics*, 5(2): 147–175.
Myers, S.C. (1984), "The Capital Structure Puzzle", *Journal of Finance*, *39*(3): 574–592.
Myers, S.C. (2001), "Capital Structure", *Journal of Economic Perspectives*, 15(2): 81–102.

Myers, S.C., and Majluf, N.S. (1984), "Corporate Financing and Investment Decisions when Firms have Information that Investors do not have", *Journal of Financial Economics*, 13(2): 187–221.

Osipov, D.V. (2012), "Capital Structure and Product Market Competition: Evidence from the EU Life Insurance Industry", Working Paper, University of Minnesota.

Ovtchinnikov, A.V. (2010), "Capital Structure Decisions: Evidence from Deregulated Industries", *Journal of Financial Economics*, 95(2): 249–274.

Pasiouras, F., and Gaganis, C. (2013), "Regulations and Soundness of Insurance Firms: International Evidence", *Journal of Business Research*, 66(5): 632–642.

Phillips, R.D., Cummins, J.D. and Allen, F., (1998), "Financial Pricing of Insurance in the Multiple-Line Insurance Company", *Journal of Risk and Insurance*, 65(4): 597–636.

Pottier, S.W., and Sommer, D.W. (1997), "Agency Theory and Life Insurer Shareholdership Structure", *Journal of Risk and Insurance*, 64(3): 529–543.

Rajan, R., and Zingales, L. (1995), "What Do We Know about Capital Structure? Some Evidence from International Data", *Journal of Finance,* 50(5): 1421–1460.

Ross, S.A. (1977), "The Determination of Financial Structure: The Incentive-Signalling Approach", *The Bell Journal of Economics*, 8(1): 23–40.

Shim, J. (2010), "Capital-Based Regulation, Portfolio Risk and Capital Determination: Empirical Evidence from the US Property–Liability Insurers", *Journal of Banking and Finance*, 34(10): 2450–2461.

Shiu, Y.M. (2011), "Reinsurance and Capital Structure: Evidence From the United Kingdom Non-Life Insurance Industry", *Journal of Risk and Insurance*,*78*(2): 475–494.

Shyam-Sunder, L., and Myers, S.C. (1999), "Testing Static Tradeoff Against Pecking Order Models of Capital Structure", *Journal of Financial Economics*, 51(2): 219–244.

Shyam-Sunder, L., and Myers, S.C. (1999), "Testing Static Tradeoff Against Pecking Order Models of Capital Structure", *Journal of Financial Economics*, 51(2): 219–244.

Slijkerman, J.F., Schoenmaker, D., and De Vries, C.G. (2013), "Systemic Risk and Diversification Across European Banks and Insurers", *Journal of Banking and Finance*, 37(3): 773–785.

Sommer, D.W. (1996). "The Impact of Firm Risk on Property-Liability Insurance Prices", *Journal of Risk and Insurance*, 63(3): 501–514.

Stiglitz, J.E. (1969), "A Re-Examination of the Modigliani-Miller Theorem", *American Economic Review,* 59(5): 784–793.

Stiglitz, J.E. (1974), "Growth with Exhaustible Natural Resources: Efficient and Optimal Growth Paths", *Review of Economic Studies*, 41: 123–137.

Titman, S., and Wessels, R. (1988), "The Determinants of Capital Structure Choice", *Journal of Finance,* 43(1): 1–19.

Van Gestel, T., Martens, D., Baesens, B., Feremans, D., Huysmans, J., and Vanthienen, J. (2007), "Forecasting and Analyzing Insurance Companies' Ratings", *International Journal of Forecasting*, 23(3): 513–529.

Weiss, M.A. (2007), "Underwriting Cycles: A Synthesis and Further Directions", *Journal of Insurance Issues*, 30(1): 31–45.

Weiss, M.A., and Chung, J.H. (2004), "US Reinsurance Prices, Financial Quality, and Global Capacity", *Journal of Risk and Insurance*, 71(3): 437–467.

Winter, R.A. (1988), "The Liability Insurance Crisis and the Dynamics of Competitive Insurance Markets", *Yale Journal of Regulation,* 5(2): 455–499.

Winter, R.A. (1994), "The Dynamics of Competitive Insurance Markets", *Journal of Financial Intermediation*, 3(4): 379–415.

Zhang, T., Cox, L.A., and Van Ness, R.A. (2009), "Adverse Selection and the Opaqueness of Insurers", *Journal of Risk and Insurance*, 76(2): 295–321.

SUBJECTIVE PERFORMANCE MEASUREMENT: A LITERATURE REVIEW

Tim Hermans*, Martine Cools and Alexandra Van den Abbeele

Abstract

This article provides an overview of academic research on subjective performance measurement, a practice that intends to remedy the weaknesses of evaluations solely based on objective quantitative performance measures. The literature on subjective performance measurement mainly focuses on four research streams: optimal contracting, discretionary bonus pools, judgment biases and debiasing, and perceived fairness. We discuss these four research streams as encountered in 67 articles published in 20 high-impact journals over the period 1977 to 2013. In addition, this article identifies several research gaps and avenues for future research.

Keywords: literature review; management control; subjective performance measurement

JEL codes:

I. INTRODUCTION

This article provides an overview of the academic literature on subjective performance measurement. This research field captures the common practice in organizations that a manager or supervisor evaluates the performance of an employee or subordinate subjectively. The subjectivity in performance evaluation can be present in several ways. Supervisors can use subjective performance measures, they can ex post adjust the weighting of objective performance measures and/or they can make discretional adjustments based on factors different from the performance measures specified ex ante (Bol, 2008; Bol & Smith, 2011). The subjective evaluations are based on personal impressions or opinions (Bol & Smith, 2011) or information not explicitly contracted on because it represents unforeseen circumstances

* Corresponding author, KU Leuven, Campus Antwerp, Korte Nieuwstraat 33, 2000 Antwerp. Email: Tim.Hermans@KULeuven.be, Tel. +32 16 37 62 50.

that would not be contractible in objective, formula-based performance evaluations (Ahn et al., 2010; Baily et al., 2011; Baiman & Rajan, 1995). Bommer et al. (1995) indicate that the correlation between objective performance measures and subjective ratings of employee performance is only 0.39. Subjective performance measures are thus clearly distinct from objective performance measures and as such sufficient attention needs to be paid to their design and use. This paper therefore offers an extensive overview of existing research on subjective performance measurement.[1]

As a research method for this literature review we searched for published articles on subjective performance measurement in the Web of Science. The following search terms were used: 'subjective performance', 'subjective evaluation', 'subjective measurement', 'subjective measure', 'subjective judgment', 'subjective assessment', 'subjective review', 'performance ratings', and 'evaluation'. We investigated whether these search terms occurred either in the topic or in the title of published articles. Publications in journals with an impact factor larger than 1 were retained for further analysis. Afterwards, we screened all obtained articles to make sure they were relevant for the purpose of this literature review.[2] This led to our final sample of 67 articles published in 20 high-impact journals over the period 1977 to 2013. We grouped these 67 articles in 4 research streams already defined in the literature based on the keywords of the most highly cited papers. For instance, Maas et al. (2012) deals with 'optimal contracting', 'discretionary bonus pools' and 'perceived fairness'. Baiman & Rajan (1995) examines 'optimal contracting' and 'discretionary bonus pools'. MacLeaod (2003) researches 'optimal contracting' and 'judgment biases and debiasing'. Baker et al. (1994), Ke et al. (1999) and Levin (2003) investigate 'optimal contracting', Gibbs et al. (2004) research 'discretionary bonus pools' and Dulebohn & Ferris (1999) and McFarlin & Sweeney (1992) look into 'perceived fairness'. Libby et al. (2004), Lipe & Salterio (2000, 2002), Moers (2005) and Prendergast & Topel (1993) focus on 'judgment biases and debiasing'. The topics of those highly cited papers resulted in 4 research streams: 'optimal contracting', 'discretionary bonus pools', 'judgment biases and debiasing' and 'perceived fairness'. Afterwards, we were able to

[1] Our study is not the first one to offer an overview of the subjective performance measurement literature. Bol (2008) examines the role of subjectivity in compensation contracts. She describes optimal contracting in a traditional agency context and thereby depicts the benefits and costs related to subjectivity in compensation contracts. Our analysis differs from the analysis of Bol (2008) because we collected a more extensive amount of papers touching more aspects of subjectivity in performance evaluation than contracting alone. Franco-Santos et al. (2012) provide a framework to classify contemporary performance measurement systems and apply this to their review of 76 empirical studies. They discuss perceptions of subjectivity, justice and trust, and judgment biases. In contrast to their general and high-level classification framework for 'all' contemporary performance measurement systems, we provide a more in-depth overview and discussion of the subjective performance measurement literature only. Prendergast & Topel (1993) discuss potential pitfalls of subjective performance evaluations: they review supervisors' preferences and biases such as leniency bias, favoritism and compression bias. We update their observations and extend the scope.

[2] Although our literature review is quite extensive, we focus on subjective performance measurement and therefore do not discuss articles dealing with feedback, performance appraisal, relative performance evaluation, self-evaluation or peer- evaluation.

fit the remaining papers of our sample in this structure based on their topic or keyword. Some papers address multiple research streams and they therefore reappear in one or more of the subsequent sections discussing each research stream separately. The remaining of this article is organized as follows. In the next sections, we discuss the four broad research streams on subjective performance measurement: optimal contracting (section 2), discretionary bonus pools (section 3), judgment biases and debiasing (section 4) and perceived fairness (section 5). Subsequently, we deal with a number of research opportunities identified through this literature review in section 6 and we end with a conclusion in section 7.

II. OPTIMAL CONTRACTING

Traditional academic research in agency theory focuses on objective performance measurement and optimal contracting. In these classical principal-agent models, a principal designs an optimal contract inducing an agent to exert effort that maximizes the value relevant to the principal. The agent gets rewarded for his effort, but he is effort-averse. The principal cannot fully observe or verify the actions undertaken by the agent and must rely on a number of objective performance measures. Appropriately designed incentive contracts can provide the agent with incentives to act in the interest of the principal and as such optimal incentive contracts mitigate or resolve agency problems (Bol, 2008; Cronqvist & Fahlenbrach, 2013).

With an optimal incentive contract, the principal does not have to monitor the agent's behavior. He can just rely on the objective outcome measures that measure the agent's performance. In other words, an agent's incentive contract provides the principal with a substitute for monitoring the agent's behavior (Morse et al., 2011). At the same time, these incentive contracts transfer risk from the principal to the agent as the objective performance measures used in these contracts do not capture the agent's effort completely and accurately. Indeed, performance in most jobs cannot be measured objectively because joint production makes individual output not readily quantifiable (Baker et al., 1988; Levin, 2003). In addition, the range of possible actions that the agent can take is too extensive to contract upon ex ante (Baker et al., 1988). As such, high-uncertainty environments warrant greater reliance on subjective performance criteria (Keeley, 1977). In practice, objective performance measures are therefore often complemented with subjective performance measures.

Table 1 provides an overview of published articles on optimal contracting including subjective performance evaluations.[3] The first article by Bol (2008) is a literature review examining the role of subjectivity in compensation contracts. In table 1 we update and extend Bol et al.'s overview. We first discuss the articles that extend the traditional agency theory

[3] The tables in this article are divided into several topics. The papers in the tables are alphabetically ordered by author name(s) within these topics.

models to account for subjectivity in performance measurement. Next we include articles revealing the benefits of subjectivity in performance contracts, to end with the articles dealing with the costs related to subjectivity in optimal contracts.

Table 1. Optimal contracting with subjective performance evaluations

Article	Focus/Results	Research Set-up
Introduction to optimal contracting		
Bol (2008)	Optimal contracting in traditional agency context: benefits and costs related to subjectivity.	Literature review
Optimal contracting models including subjectivity		
Baiman & Rajan (1995)	Discretionary bonus pools are an efficient way to incorporate non-contractible information in a two-agent setting.	Analytical model
Budde (2007)	Researches a BSC with contractible and not-contractible scorecard measures: a combination of a formal contract and a subjective performance evaluation may outperform a purely formal contract.	Analytical model
Cronqvist & Fahlenbrach (2013)	Private equity sponsors (strong principals) use less subjective performance measures, but some subjective performance evaluation to compensate the CEO.	Field study (CEO contract data on leveraged buyouts of 20 large, American listed firms, 2005–2007)
Höppe & Moers (2011)	Different types of subjectivity are used for different purposes: "discretionary bonuses" are used for risk-reduction, "subjective weights" for congruity-improvement.	Archival study (1,753 firm-year- observations for 424 American, publicly listed firms, 1998–2002)
Ke, Petroni & Safieddine (1999)	Privately held insurers (strong principals) use more subjective performance measures to compensate the CEO.	Archival data (45 privately-held and 18 publicly-held American insurers, 1994–1996)
MacLeod (2003)	Extends standard principal-agent model with a single agent with subjective evaluations.	Analytical model
Rajan & Reichelstein (2006)	When the bonus pool covers many agents and/or the principal's subjective information is precise, discretionary bonus pools are nearly as efficient as explicit contracts.	Analytical model
Rajan & Reichelstein (2009)	In the single-agent case it might be optimal to ignore the subjective signal with discretionary bonus pools.	Analytical model
Benefits of subjectivity in optimal contracts		
Baker, Gibbons & Murphy (1994)	A combination of objective and subjective measures sometime outperforms an explicit or an implicit contract alone.	Analytical model
Baker, Jensen & Murphy (1988)	Discusses several benefits and costs related to objective and subjective performance measurement.	Literature review

Article	Focus/Results	Research Set-up
Gibbs, Merchant, Van der Stede & Vargus (2004)	Subjective bonuses are used to complement perceived weaknesses in quantitative performance measures and to provide employees insurance against downside risk in their pay.	Archival study (526 department managers in 250 American car dealerships in 1998–1999) and 1050 surveys in 326 different dealerships
Höppe & Moers (2011)	"discretionary bonuses" are used for risk-reduction, "subjective weights" for congruity-improvement.	Archival study (1,753 firm-year-observations for 424 American, publicly listed firms, 1998–2002)
Indjejikian & Matejka (2012)	Nonfinancial measures or subjective evaluations are more used for bonuses when the recipients have greater influence over the internal accounting systems design.	Survey (242 BU-managers and controllers of 121 Bus of 7 Dutch multinationals and 48 additional interviews)
Ke, Petroni & Safieddine (1999)	Privately held insurers (strong principals) use more subjective performance measures to compensate the CEO.	Archival data (45 privately-held and 18 publicly-held American insurers, 1994–1996)
Keeley (1977)	High-uncertainty environments warrant greater reliance on subjective performance criteria.	Questionnaire (106 supervisor-subordinate pairs)
Costs of subjectivity for optimal contracts		
Ahn, Hwang & Kim (2010)	Subjective measures provide less incentive than objective measures because of the lack of variation in scores (compression bias).	Archival (13 government-invested companies, Republic of Korea, 1990–2006)
Baker, Jensen & Murphy (1988)	Discusses several benefits and costs related to objective and subjective performance measurement.	Literature review
Golman & Bhatia (2012)	Subjective performance evaluation leads to leniency bias, and associated with that reduced employee effort.	Analytical model
Krishnan, Luft & Shields (2005)	Individuals do not weigh measures appropriately in a two-measure incentive system.	Experiment (32 accounting and MBA students)
Levin (2003)	Self-enforced relational contracts with moral hazard result in compression bias.	Analytical model
MacLeod (2003)	Optimal contracts with subjective evaluations can result in compression bias and leniency.	Analytical model
Morse, Nanda & Seru (2011)	Powerful CEOs can shift the weight on performance measures toward the better performing measures. This affects future firm performance negatively.	Archival study with 2348 firms over the period 1992–2003
Prendergast & Topel (1996)	Subjectivity leads to favoritism	Analytical model
Prendergast (1993)	With subjective evaluation procedures, subordinates have an incentive to conform to what they feel their superiors want to hear.	Analytical model

The first eight articles extend traditional agency theory knowledge with one or several aspects of subjectivity. MacLeod (2003) allows for subjective performance evaluations in the standard principal-agent model by including subjective performance measures. With this analytical model, he shows that if the principal's and the agent's subjective evaluations correspond (or equivalently if there is trust and perceived fairness between principal and agent), one can implement the optimal contract just as if subjective evaluations were objective and verifiable. Budde (2007) provides a theoretical model for a combination of objective and subjective performance measures in a balanced scorecard (BSC) setting. The model shows that when all objective performance measures are perfectly verifiable, a properly designed BSC can perfectly align the interests of the principal and the agents with an explicit contract.

When not all BSC measures are contractible[4], the first-best solution, a contract in which the agent exerts the optimal level of effort that provides the optimal value relevant to the principal, may still be obtained through a combination of a formal contract and a subjective performance evaluation (Budde, 2007). Höppe & Moers (2011) undertook an archival study in which they focus on the use of two different types of subjectivity: "subjective weights" and "discretionary bonuses". "Subjective weights" concern the option whereby supervisors can ex post adjust the weighting of objective performance measures, while "discretionary bonuses" refer to the case where supervisors can make discretional adjustments based on factors different from the performance measures specified ex ante. According to optimal contracting considerations, their results show that "subjective weights" are used to improve goal congruence between the agent and the principal, while "discretionary bonuses" are used to reduce risk for the agent due to uncertainty.

Baiman & Rajan (1995) and Rajan & Reichelstein (2006, 2009) provide analytical models on the use of discretionary bonus pools. For a discretionary bonus pool, the bonus pool is based on an explicit formula involving objective performance measures agreed-upon ex ante. Afterwards, the bonus pool is allocated amongst the agents at the principal's discretion. The entire bonus pool is paid out regardless of the subjective information observed by the principal, but in case of unfavorable subjective information the principal withholds part of the bonus of one agent to give it to other, better-performing agents. Baiman & Rajan (1995) prove that discretionary bonus pools result in a strict Pareto improvement compared to the optimal contract that does not use non-contractible information by enabling a principal to exploit non- contractible information to motivate agents. Furthermore, Rajan & Reichelstein (2006) show that discretionary bonus pools are optimal when a principal must rely solely on non-verifiable, subjective information to create incentives for a group of agents. They find that bonus pools are nearly as efficient as explicit contracts, provided that the bonus pool covers a large number of agents and/or the principal's subjective information is fairly precise.

[4] A performance measure is contractible if its value is observable both by the principal, the agent and an unrelated third party. In this case the performance measure can be explicitly incorporated in a contract (Baiman & Rajan, 1995).

In addition, when no other agent is present, the principal incurs an additional cost when the agent shirks. The model of Rajan & Reichelstein (2009) indicates that in the single-agent case it might be optimal to ignore the subjective signal. When both objective and subjective measures are used, the optimal contract results in less divergent performance scores relative to the number of performance levels on the different performance measures than when only objective measures are used. Furthermore, they show that the single-agent bonus pool results in less divergent performance scores relative to the number of possible performance scores on the different performance measures than a multiple-agent bonus pool (Rajan & Reichelstein, 2009).

In agency models extended with subjective performance measures, strong principals have greater incentives to observe and monitor agents' effort and to base agents' reward on those subjective observations. Consequently, agent's compensation is less likely based on an explicit contract with objective performance measures (Ke et al., 1999). Ke et al. (1999) confirm this theoretic reasoning empirically via an archival study amongst privately-held and publicly-held property-liability insurers. They find that within privately-held insurers (called strong principals) CEO compensation is less based on objective measures like accounting information and presumably more on subjective measures compared to the publicly-held insurers (called weak principals). Consequently, their findings are consistent with optimal contracting (Ke et al., 1999). Cronqvist & Fahlenbrach (2013) come to the opposite conclusion in their study of CEO contracts within large American firms moving from public ownership with dispersed owners (weak principal) to private ownership with strong principals. They find that strong principals redesign contracts away from qualitative, nonfinancial measures, but they introduce subjective performance evaluation instead. Baker et al. (1994) assume objective performance measures are imperfect and cause incentive distortions, which can be mitigated by the inclusion of subjective performance assessments. The authors prove that in some circumstances, neither an explicit nor an implicit contract alone yields positive profit, but a combination of objective and subjective measures can.

Next, we discuss seven articles on the benefits of subjectivity in optimal contracts. Gibbs et al. (2004) use archival and survey data on compensation of managers in car dealerships to examine when firms make greater use of subjectivity in bonus payments. It turns out that subjective bonuses are used to respond to perceived weaknesses in quantitative formulaic bonuses such as incompleteness, short-term focus and susceptibility to manipulation. Using – only imperfect – objective performance measures may lead to suboptimal actions taken by the agents (Baker et al., 1994). Agents tend to focus their effort on the directly rewarded activities and away from the unrewarded activities. The misspecification of an objective performance measurement system thus may result in agents "gaming the system" by optimizing actual instead of intended measures (Baker et al., 1988). As such, contracts based solely on objective performance measures are imperfect and cause incentive distortions. This problem can be mitigated by including (additional) subjective performance assessments (Baker et al., 1994; Höppe & Moers, 2011). Indjejikian & Matejka (2012) involved business

unit managers and controllers in a survey study supplemented with in-depth interviews. They find that principals rely more on nonfinancial measures or subjective evaluations in determining local managers' bonuses when local managers have a greater influence on the design of internal accounting systems. This is consistent with principals protecting themselves against agents' asymmetric information or agents' manipulation of objective, accounting measures. Baker et al. (1994) theorize that in some circumstances a combination of objective and subjective measures outperforms an explicit or an implicit contract alone. Moreover, the subjective bonuses provide employees insurance against downside risk in their pay e.g. by filtering out the effect of uncontrollable factors due to interdependencies (Gibbs et al., 2004) or uncertainty (Keeley, 1977; Höppe & Moers, 2011), recalculating incentives when performance targets are too challenging or when the department is facing losses. Subjectivity improves incentive contracting when there is greater trust between the subordinate and the supervisor. This is because the positive effects of subjective bonuses on pay satisfaction and firm performance are larger the longer the supervisor's tenure due to mutual trust (Gibbs et al. 2004). In sum, principals use subjectivity to resolve contracting problems such as incentive distortions (congruity issues) (Höppe & Moers, 2011; Baker et al., 1994), risk concerns (Höppe & Moers, 2011), environmental uncertainty (Keeley, 1977), moral hazard (Cronqvist & Fahlenbrach, 2013; Ke et al., 1999), asymmetric information or agents "gaming" (Indjejikian & Matejka, 2012) or manipulating the system (Gibbs et al.; 2004).

However, allowing subjectivity in performance evaluations also has its downsides, as described in the nine articles discussed next. Levin (2003) argues that the use of subjective performance measures necessarily leads to costly disputes and conflicts between the agent and the principal. When agents feel their evaluation is unfair, fairness and conflict concerns will lead to compressed and above average subjective performance evaluations (and thus to higher compensation for the agent) (Golman & Bhatia, 2012; Levin, 2003; MacLeod, 2003). This compressed and above average rating behavior in turn lowers employee performance and firm productivity (Ahn et al., 2010; Golman & Bhatia, 2012). Another concern related to (un)fair rating behavior is favoritism. Subjectivity leads to favoritism where evaluators act on personal preferences toward subordinates to favor some employees over others beyond their true performance. This reduces incentives for the other agents because of increased risk/uncertainty in evaluations (Prendergast & Topel, 1996). In addition, individuals seem insufficiently aware that a change in the accounting for one subjective measure has spillover effects on the optimal weighting of the other subjective measure in a two-measure incentive system. Consequently, they make performance-measure weighting decisions that are likely to result in misallocations of agent effort (Krishnan et al., 2005). Morse et al. (2011) provide archival evidence that powerful agents are able to shift the weight on performance measures toward the better performing measures. This manipulation practice harms future firm performance. In addition, Prendergast (1993) theorizes that agents have an incentive to conform to what they feel their superior wants to hear. The agent distorts his opinion towards the anticipated opinion of the supervisor. As such, too much weight is put on the opinion of

the supervisor, which leads to inefficiencies. Another difficulty in subjective performance measurement is due to principals reneging, i.e. they assess the agent's final performance untruthfully in order to pay less reward to the agent. This is possible because the subjective performance information in the optimal contract is not enforceable (Ahn et al., 2010; Baily et al., 2011; Baiman & Rajan, 1995; Baker et al., 1994; Bol, 2008; Bol & Smith, 2011; MacLeod, 2003; Prendergast & Topel, 1993). This evidently undermines the credibility of subjective performance evaluations in optimal contracts. Therefore, it is important to install mechanisms to enforce the subjective performance measurement. Baker et al. (1994) for example cite that implicit contracts are self-enforcing as principals are concerned with their reputation in the labor market for keeping their promises. Levin (2003) remarks that each party has the option to walk away in a relational contract. To prevent that the principal reneges, the payable reward must not exceed the net present value of the benefits the principal realizes under an ongoing contract. This is the case if the principal's discount rate is small enough. The credibility of optimal contracts with subjectivity added can be considerably improved by restricting subjective incentives to that part of the first-best action that cannot be induced by an explicit contract (Budde, 2007). In addition, "discretionary bonus pools" could prevent the principal from reneging in a situation with multiple agents, because the bonus pool amount is agreed upon ex ante and afterwards the total bonus pool is allocated amongst the agents according to the principal's discretion (Baiman & Rajan, 1995; Rajan & Reichelstein, 2006, 2009).

III. DISCRETIONARY BONUS POOLS

In this section, we first discuss a number of theoretical articles on the characteristics and benefits of discretionary bonus pools, after which we discuss the experimental articles challenging the predictions in the theoretical articles. Next, we discuss articles investigating the use of discretionary bonus pools in practice. Table 2 follows the structure of this section.

Table 2. Discretionary bonus pools

Article	Focus/Results	Research Set-up
Analytical models on discretionary bonus pools		
Baiman & Rajan (1995)	Discretionary bonus pools are an efficient way to incorporate non-contractible information in a two-agent setting.	Analytical model
Baker, Jensen & Murphy (1988)	Free-rider problem associated with ordinary profit-sharing plans: employees receive only approximately 1/n of the increased profits related to their effort (where n is the number of participants in the plan).	Literature review

Article	Focus/Results	Research Set-up
Rajan & Reichelstein (2006)	When the bonus pool covers many agents and/or the principal's subjective information is precise, discretionary bonus pools are nearly as efficient as explicit contracts.	Analytical model
Rajan & Reichelstein (2009)	In the single-agent case it might be optimal to ignore the subjective signal with discretionary bonus pools.	Analytical model
Experimental papers on discretionary bonus pools		
Bailey, Hecht & Towry (2011)	Managers incorporate noncontractible information to a lesser extent than theoretically expected. Managers who can only allocate a part of the bonus pool incorporate noncontractible information to a greater extent than participants with full discretion.	Experiment (170 business school students)
Fisher, Maines, Peffer & Sprinkle (2005)	Subordinate's performance and compensation is larger when the employer has no discretion over total employee compensation, but discretion over allocation of the compensation pool.	Experiment (237 undergraduate business students)
Maas, van Rinsum & Towry (2012)	Supervisors are more willing to obtain costly performance information on individual agents as it becomes more difficult to distinguish individual contributions to group performance.	Experiment (126 undergraduate business students)
Field studies on discretionary bonus pools		
Gibbs, Merchant, Van der Stede & Vargus (2004)	Discretionary bonuses are used to complement perceived weaknesses in quantitative performance measures and to provide employees insurance against downside risk in their pay.	Archival study (526 department managers in 250 American car dealerships in 1998–1999) and 1050 surveys in 326 different dealerships
Ittner, Larcker & Meyer (2003)	Discretion in weighting the measures in a BSC bonus plan led to a focus on quantitative, outcome-oriented financial performance measures that were used in earlier non-discretionary bonus plans.	Field study (a large American retail bank)
Ivancevich (1983)	The more unsatisfactory performing engineers in a team, the more favorable ratings are for satisfactory performing engineers. For scientists no such effect was identified.	Field study with 104 supervisors of 624 engineers and 66 supervisors of 404 scientists working in an American company
Merchant, Chow & Wu (1995)	Incentive plans in Taiwanese and US firms are very similar. In both countries, firms make use of discretionary bonus pools.	Field study (open-ended interviews in 2 US and 2 Taiwanese companies)

Baiman & Rajan (1995) and Rajan & Reichelstein (2006, 2009) theoretically discuss the characteristics and benefits of discretionary bonus pools. As indicated above, subjective performance information is complex and subtle, and therefore difficult to observe and verify by a third party. Since this information is not enforceable (Ahn et al., 2010; Baily et al., 2011; Baiman & Rajan, 1995; Baker et al., 1994; Bol, 2008; Bol & Smith, 2011; MacLeod, 2003; Prendergast & Topel, 1993), it harms the credibility of subjective performance measurement in optimal contracts. Supervisors can assess a subordinate's final performance untruthfully in order to pay less reward to that subordinate (Bol, 2008). In this context, a discretionary bonus pool is an appealing instrument. The magnitude of the bonus pool is based on an explicit formula agreed-upon ex ante and involving objective performance measures. The entire bonus pool is paid out regardless of the subjective information observed by the supervisor (Baiman & Rajan, 1995). Supervisors cannot change the magnitude of the reward by assessing agent's performance falsely and they have consequently no incentive to do so (Rajan & Reichelstein, 2006). Afterwards, the bonus pool is allocated amongst different subordinates at the supervisor's discretion. Based on subjective information the supervisor can shift a part of the bonus of one subordinate to another, better-performing colleague. Accordingly, a supervisor can use non-contractible information to encourage subordinates (Baiman & Rajan, 1995). The supervisor discretion solves the free-rider problem associated with ordinary profit-sharing plans in large organizations described by Baker et al. (1988). With ordinary profit-sharing plans, employees bear the full cost of exerting effort and yet receive only 1/n of the increased profits (where n is the number of participants in the plan). Discretionary bonus pools take individual effort into account (Baiman & Rajan, 1995; Rajan & Reichelstein, 2006, 2009). The use of non-contractible information to motivate subordinates results in a strict Pareto improvement compared to the optimal contract that does not use non-contractible information (Baiman & Rajan, 1995). Discretionary bonus pools are even optimal if a supervisor can only rely on non-verifiable subjective information to create incentives for a group of subordinates. Provided the bonus pool covers a large number of subordinates and/or the supervisors' subjective information is fairly precise, bonus pools based solely on subjective information should be nearly as efficient as explicit contracts based on objective and verifiable information (Rajan & Reichelstein, 2006). When no other subordinate is present, the supervisor incurs an additional cost when the subordinate shirks. In the single-subordinate case it might be optimal to ignore the subjective signal (Rajan & Reichelstein, 2009). The reasoning above explains the popularity of discretionary bonus pools both in practice and in research.

Three experimental articles challenge the predictions made by the theoretical articles discussed above. Fisher et al. (2005) undertook an experiment to examine situations in which the supervisor either has full discretion or no discretion over the *magnitude* of the bonus pool and/or the *allocation* of this bonus pool among subordinates. A compensation scheme in which a supervisor has full discretion to use private information may reduce subordinate opportunism, but allows for supervisor opportunism. The researchers measured the total group output of the subordinates, the bonus allocated to the subordinates and the residual

supervisor profit. Both total group output and subordinate compensation appeared to be greater when the supervisor had no discretion over the magnitude of the bonus pool, but discretion over the allocation of the bonus pool. The supervisor's residual profit was higher when he had discretion over the allocation of compensation, while discretion over the magnitude of the bonus pool had no effect on residual profit. So, in general, the discretionary bonus pool outperforms the other experimental situations. This is consistent with Baiman & Rajan (1995). Bailey et al. (2011) experimentally examine situations in which the supervisor has *full or partial* discretion to allocate the bonus pool and/or he is confronted with *positive or negative* noncontractible information. The findings show that managers incorporate noncontractible information to a lesser extent than theoretically expected by Rajan & Reichelstein (2006) when allocating a bonus pool. When processing performance information, managers in the experiment tended to choose an anchor point and then subsequently adjusted for noncontractible information. This anchoring approach is in contrast to the theoretical approach in Rajan & Reichelstein (2006) in which a manager is supposed to integrate all contractible and noncontractible information into a single, comprehensive performance measure (integrative approach). Managers who use an anchoring approach incorporate noncontractible information into bonus pool allocations to a lesser extent than those who use an integrative approach. In practice, this leads to a reduction in the intended, theoretical benefits of managerial discretion in bonus allocation proclaimed by Rajan & Reichelstein (2006). Participants who can only allocate a part of the bonus pool incorporate noncontractible information to a greater extent than participants with full discretion (Bailey et al., 2011). The third experimental article, by Maas et al., starts from the observation that joint production and unobservability make individual output not readily quantifiable in most jobs (Baker et al., 1988; Levin, 2003; Maas et al., 2012). This impedes the bonus pool allocation discretion of the supervisor in a discretionary bonus pool setting. Maas et al. (2012) investigate the willingness of supervisors to obtain additional, costly information to more accurately assess individual contributions to team output. In their experiment, the aggregate team output is readily available and the individual output can be obtained at an additional cost. The results indicate that supervisors are willing to incur a cost to prevent potential unfairness. Supervisors are more willing to obtain the costly information as it becomes more difficult to distinguish individual contributions to group performance. Additionally, this willingness appeared to be greater for relatively high versus relatively low levels of group performance.

Four articles investigate the functioning of discretionary bonus pools in practice. Ivancevich (1983) provides evidence that a supervisor shifting a part of the bonus from one subordinate to another, better-performing colleague (supervisor allocation discretion) follows a natural reflex. In a field study, Ivancevich (1983) instructed supervisors to evaluate each member of their team individually. Team size ranged from 9 to 44 engineers per team. The supervisors were instructed not to force themselves to come up with distributed evaluations. Despite this instruction, the study shows contrast effects that are very similar to situations where supervisors use allocation discretion in discretionary bonus pools. The

more unsatisfactory performing engineers in a team, the more favorable the ratings are for the well performing engineers. Well performing employees are thus the beneficiaries of higher performance ratings and more rewards when unsatisfactory performers are part of the supervisor's team (Ivancevich, 1983). Merchant et al. (1995) investigate the use of discretionary bonus pools in practice. They show that incentive plans in Taiwanese and US firms are very similar in terms of the extent of individual performance-dependent monetary rewards, the extent of group-rewards compared to individual rewards and the amount of subjectivity in evaluations.[5]

Subsequently, Gibbs et al. (2004) find in the car dealership context that discretionary bonuses are used to complement perceived weaknesses in quantitative performance measures (incompleteness, short-term focus and susceptibility to manipulation) and to provide employees with an insurance against downside risk in their pay (by filtering out uncontrollables due to interdependencies, recalculating incentives when performance targets are too challenging or when department is facing losses). In addition, they find that the use of discretionary bonus pools is positively related to pay satisfaction and firm performance when the manager has long tenure. Finally, Ittner et al. (2003) undertake a field study on the introduction of a subjective BSC-based bonus plan containing six categories of financial and nonfinancial performance measures in a large American retail bank. The supervisor could subjectively decide on the weighting of the different performance measures. Ittner et al. (2003) were confronted with a number of downsides of this practice. The discretion in weighting the measures in the bonus plan led the supervisors to ignore many performance measures, to change weightings from period to period and to include factors that were not even performance measures, although this was not allowed. In other words, quantitative, outcome-oriented financial performance measures as used in earlier non-discretionary bonus plans remain dominant. The high level of discretion and the related uncertainty in the criteria used for bonus determination made many subordinates complain about favoritism. Afterwards, the firm chose for a non-discretionary, formulaic bonus plan based solely on revenues. This field study points out that psychology may be more important in explaining firm's measurement practices than optimal contracting.

IV. JUDGMENT BIASES & DEBIASING

Rating inaccuracy caused by performance evaluation biases is perceived as one of the main problems of introducing subjectivity into compensation contracts (Bol, 2011). Supervisors need to invest time and effort in gathering accurate information on employee performance (Bol, 2011) and are not the residual claimants of subordinates' output, which leaves room for supervisors' preferences (Prendergast & Topel, 1993). This section provides an overview of

[5] While no cultural differences were found, we should be cautious when interpreting these results as the researchers only investigated 2 US companies and 2 Taiwanese companies (Merchant et al., 1995).

different judgment biases by supervisors. These judgment biases may impede or reinforce the proclaimed benefits of subjectivity in performance measurement discussed in the previous section on optimal contracting. Table 3 provides an overview of the articles on judgment biases and debiasing. Debiasing concerns practices to resolve judgment biases. It lists the articles discussing compression bias or centrality bias, the articles focusing on biases related to the BSC, the articles concerning biases related to personal characteristics, and a number of articles about biases related to accompanying or competitive information.

Table 3. Judgment biases and debiasing in subjective performance evaluations

Article	Focus/Results	Research Set-up
Leniency bias and compression bias		
Ahn, Hwang & Kim (2010)	Subjective measures provide less incentive than objective measures because of the lack of variation in scores (compression bias).	Archival (13 government-invested companies, Republic of Korea, 1990–2006)
Baker, Jensen & Murphy (1988)	Biased and inaccurate performance evaluations reduce effectiveness of incentives and productivity.	Literature review
Bol (2011)	Information-gathering costs and strong subordinate-supervisor relationships increase centrality bias and leniency bias. Centrality bias decreases performance improvement and leniency bias increases future performance.	Archival study (5 branch offices of a Dutch financial service provider, 2003–2004, 198 employees
Duarte, Goodson & Klich (1994)	Subjective performance ratings in high-quality supervisor-subordinate relationships are always high. Ratings in low-quality relationships are consistent with objective performance measures in the short run, but high in the long run.	Questionnaire (261 supervisor- subordinate pairs in an American telephone company)
Golman & Bhatia (2012)	Noise in the performance signal and a stronger aversion to unfairly low ratings than to overly high ones result in leniency bias. Noise in the performance signal results in compression bias. Both biases hurt agent's performance.	Analytical model
Judge & Ferris (1993)	Greater opportunity to observe a subordinate's job performance resulted in higher performance ratings.	Questionnaire (81 nurses and their supervisors in an American Hospital)
Kane, Bernardin, Villanova & Peyrefitte (1995)	Leniency is a relatively stable response tendency by individual raters.	3 field studies: One: 328 patrol officers, 38 sergeants and 14 lieutenants in a police department, two: 243 nurses and 31 head nurses, three: 44 supervisors of 376 social workers

Article	Focus/Results	Research Set-up
Levin (2003)	Self-enforced relational contracts with moral hazard result in compression bias.	Analytical model
MacLeod (2003)	Optimal contracts with subjective evaluations can result in compression bias and leniency bias.	Analytical model
Moers (2005)	The use of multiple objective performance measures and the use of subjective performance measures are related to compression and leniency bias.	Archival study (124 subordinates in a Dutch, maritime industrial firm, 1998)
Prendergast & Topel (1993)	Supervisors are not the residual claimants of subordinates' output, which makes supervisors' preferences and biases such as leniency bias, favoritism and compression bias possible.	Literature review
Biases observed in a BSC context		
Banker, Chang & Pizzini (2004)	Evaluators focus more on common measures than on unique measures. Evaluators focus more on strategically linked measures than non-linked measures only when evaluators are provided with detailed information about BU-strategy.	Experiment (480 MBA students)
Cardinaels & van Veen-Dirks (2010)	When there are performance differences in the financial performance measures, evaluators that use a BSC- format place more weight on the financial performance measures than evaluators using an unformatted scorecard. When there are performance differences in the non-financial performance measures, evaluators evaluate similarly in both formats.	2 experiments (144 business program students)
Choi, Hecht & Tayler (2012)	Surrogation: managers forget that performance measures are imperfect representations of the underlying strategic construct.	Experiment (79 graduate business students)
Ding & Beaulieu (2011)	Participants who were induced to feel good (bad) gave higher (lower) evaluation scores to divisional managers.	Experiment 1 (104 MBA students) and Experiment 2 (32 MBA students)
Humphreys & Trotman (2011)	When *all* the performance measures are strategically linked, but no strategy information is provided, common measure bias exists. When strategy information is present and *all* performance measures are strategically linked, then common measure bias disappears.	Experiment (92 executive MBA students)
Libby, Salterio & Webb (2004)	Either the requirement to justify a performance evaluation to a superior or improving perceived quality of the BSC measures via an independent third-party report decreases common measure bias.	Experiment (227 MBA students)

Article	Focus/Results	Research Set-up
Lipe & Salterio (2000)	Superiors use only the common performance measures to evaluate performance of the business unit in a BSC- context.	Experiment (58 MBA students)
Lipe & Salterio (2002)	Performance evaluations are affected by organizing the measures into the BSC categories when multiple below- target (or above-target) measures are contained within a category but those evaluations are not affected when the above/below-target measures are distributed across the scorecard's four categories or when the same measures are presented without the organizing BSC categories.	Experiment (78 MBA students)
Tayler (2010)	Managers who are involved in selecting strategic initiatives perceive those initiatives afterwards as more successful than managers who are not involved in the initiative-selection process. Simply framing the scorecard as a causal chain is not sufficient to mitigate these effects, but framing the scorecard as a causal chain *and* involving managers in the selection of scorecard measures, mitigates the effects.	Experiment (132 MBA students)
Biases related to personal characteristics		
Biernat & Sesko (2013)	Evaluations of mixed-sex work teams' performance: women were solely judged lower in a white pair work team. Black women were not affected by gender bias.	2 experiments (142 and 283 undergraduate students respectively)
Elvira & Town (2001)	Racial differences between subordinate and supervisor lead to lower ratings for both black and white subordinates.	Field study (316 salespersons in a large, American company)
Judge & Ferris (1993)	Demographic similarity and the supervisor-subordinate relationship significantly influenced performance rating.	Questionnaire (81 nurses and their supervisors in an American Hospital)
Pulakos & Wexley (1983)	Supervisors appraise dissimilar subordinates significantly lower.	Questionnaire (171 supervisor- subordinate relationships in manufacturing, retailing, government and service organizations)
Varma & Stroh (2001)	Both male and female supervisors rate subordinates of the same sex higher.	220 surveys of supervisors in the communications industry
Wayne & Liden (1995)	Demographic similarity and subordinates' impression management influence performance ratings.	Survey (111 supervisor-subordinate pairs in nonacademic jobs at 2 American universities)

Article	Focus/Results	Research Set-up
Biases related to accompanying or competitive information		
Bol & Smith (2011)	Supervisors bias their subjective evaluations of performance to be consistent with an accompanying objective performance measure.	Experiment (216 supervising employees at an university)
Dossett & Greenberg (1981)	A worker who initially suggested a high goal received a significantly higher performance score than a worker who suggested a low goal.	Experiment with 80 undergraduate students
Ghosh & Lusch (2000)	Unfavorable outcomes negatively influence subjective performance evaluations.	Archival study in 204 stores of an American retailer
Heneman & Wexley (1983)	Performance ratings are less accurate when rating is delayed and when only a small amount of information is observed.	Experiment (180 undergraduate business students)
Hogan (1987)	Ratings will be lower when a subordinate's actual performance disappoints a supervisor's expectations about that performance.	Questionnaire (49 subordinate- supervisor pairs in an American bank)
Ittner, Larcker & Meyer (2003)	Supervisors focus on quantitative, outcome-oriented financial performance measures.	Field study (a large American retail bank)
Jacobs & Kozlowski (1985)	As raters have more opportunity to observe ratee behavior, the magnitude of halo error increases.	3 consecutive ratings (1031, 976 and 876 students respectively).
Tan & Jamal (2001)	Average superiors evaluate work done by outstanding subordinates more favorably than work done by average subordinates when they know the identity of the work preparer. Outstanding superiors are not affected by the halo effect.	Experiment (40 audit seniors and 20 audit managers)

Compression bias or centrality bias refers to the tendency to compress performance ratings, which results in less variance in ratings than justified by the variance in actual performance. Leniency bias is the tendency to inflate subordinate's performance rating such that subordinate's performance is assessed to be above average (Baker et al., 1988; Bol, 2011). These important forms of performance evaluation bias have received quite some research attention. According to the theoretical articles on optimal contracting, supervisor's subjective performance evaluations will be compressed and rated above average if the supervisor and subordinate disagree upon the subjective performance evaluation or when moral hazard is present (Levin, 2003; MacLeod, 2003). This is due to the tradeoff between reducing the cost of conflict between subordinate and supervisor ex post (Bol, 2011; MacLeod, 2003) and providing incentives to the subordinate ex ante (MacLeod, 2003). Supervisors need to invest time and effort in gathering accurate information on employee performance (Bol, 2011). They bear all of the monitoring costs but receive little of the benefit

from conducting more accurate evaluations (Baker et al., 1988). Whenever information-gathering costs increase, they invest less time and effort in gathering accurate information on employee performance. Therefore, it is not surprising that empirical evidence indicates that centrality bias and leniency bias are positively related to information-gathering costs (Bol, 2011), to the use of multiple objective performance measures and to the use of subjective performance measures (Moers, 2005). Additionally, uncertainty about subordinate performance leads to compressed ratings. The analytical model by Golman & Bhatia (2012) indicates that when a supervisor is uncertain about subordinate performance and he has a stronger aversion to unfairly low ratings than to overly high ratings (due to fairness or conflict concerns), he will inflate performance ratings according to his preferences. Kane et al. (1995) provide empirical evidence that inflating performance ratings is a relatively stable rater tendency.

In contrast to the reasoning above, Duarte et al. (1994) and Judge & Ferris (1993) provide empirical evidence about the presence of leniency in a different way: they found that a greater opportunity to observe a subordinate's job performance actually resulted in higher performance ratings. Along this reasoning, the supervisor-subordinate relationship significantly influences supervisors' affection toward subordinates and consequently influences performance ratings indirectly through supervisors' affection (Judge & Ferris, 1993). Strong subordinate-supervisor relationships increase centrality bias and leniency bias (Bol, 2011). Both in the short run and the long run, subjective performance ratings in high-quality supervisor-subordinate relationships are high, regardless of objective performance measures. Ratings in low-quality relationships are consistent with objective performance measures in the short run, but high in the long run, regardless of objective measures (Duarte et al., 1994).

The literature shows mixed results on the effect of centrality bias and leniency bias on subordinate performance. Ahn et al. (2010) examine the effect of discriminability (variation in performance scores) on subordinate performance empirically. Their findings show that subordinate performance improvement increases with the degree of discriminability. Subjective measures provide less incentive than objective measures because of the lack of discriminability (compression bias). This results in a decrease in performance improvement (Ahn et al., 2010; Bol, 2011). Biased and inaccurate performance evaluations reduce productivity by reducing effectiveness of incentives (Baker et al., 1988). In other words, leniency bias and centrality bias hurt the agent's performance (Golman & Bhatia, 2012). In contrast, Bol (2011) reveals that leniency bias increases future performance due to increased perceived fairness of the incentive system.

In the following paragraphs, we discuss a number of experiments investigating judgment biases in the context of the BSC or another strategic performance evaluation framework. The BSC is a framework devised by Kaplan & Norton (1992) containing a large set of performance measures that capture the drivers of the firm's desired business strategy along four categories: financial performance, customer relations, internal business processes and the organization's

learning and growth activities. The distinctive feature of the BSC is that the performance measures are linked with each other, and that they express cause-and-effect relationships that lead to the implementation of the intended strategy. The BSC can be used to evaluate the performance of a business unit or a business unit manager (Lipe & Salterio, 2000, 2002). Lipe & Salterio (2000) point out that BSCs include some performance measures common to multiple business units and other performance measures that are unique to a particular business unit. Based on an experiment, they discover that supervisors make only use of the common measures when evaluating the performance of different business units i.e. common measure bias is present (Lipe & Salterio, 2000). Banker et al. (2004) confirm this result in an experiment in which some measures are strategically linked and others are not: evaluators focus more on common performance measures than on unique measures. However, this result does not hold when a number of the BSC performance measures are strategically linked *and* detailed information on the strategic linkages is provided. In that case, evaluators focus more on strategically linked unique measures than on non-linked measures that are common (Banker et al., 2004). In contrast, when strategy information is provided to managers *and* only *some* measures are strategically linked, common linked measures get more attention than unique linked measures (Banker et al., 2004; Humphreys & Trotman, 2011). Humphreys & Trotman (2011) further experimentally demonstrate that common measure bias exists when only some or *all* the performance measures are strategically linked, but no strategy information is provided. However, when strategy information is present and *all* performance measures are strategically linked, common measure bias disappears, but Libby et al. (2004) propose two methods to overcome common measure bias. The first one is to introduce a requirement to justify the performance evaluation to a superior. An alternative is to improve the perceived quality of the BSC performance measures via the provision of an independent third-party report (Libby et al., 2004).

Other experimental studies focus on how BSC framing affects performance evaluations. Lipe & Salterio (2002) find that when multiple below-target (or above-target) measures are contained within a single BSC category, performance evaluations are different from performance evaluations using the same measures, but without the BSC categories framework. However, when the above/below-target performance measures are distributed across the four categories of the BSC, evaluations are not different from evaluations using these same measures, but without the BSC categories framework. The reason is that when performance on measures within a group is consistent (e.g. consistently above-target), the decision maker may perceive that the measures are related. Consequently, he reduces the impact of the individual performance measures on his or her judgment. In contrast, when the same measures are presented without the organizing BSC categories (or are scattered across BSC categories), the perception of relations among these measures and the resulting reduction in decision weights are less likely. Also Cardinaels & van Veen-Dirks (2010) investigate the effect of the presentation of performance measures on performance evaluations of two business unit managers, especially how evaluators weight financial and non-financial

measures. When the performance difference between the two managers is located in performance measures in the financial category, evaluators that use a BSC-format place more weight on the financial category performance measures than evaluators using an unformatted scorecard. In contrast, when the performance difference is located in performance measures of one of the three non-financial categories, the weight placed upon these non-financial category measures is similar for the BSC-format and the unformatted scorecard. In a subsequent experiment, Cardinaels & van Veen-Dirks (2010) use performance markers to indicate above- target, on-target or below-target performance. In this setting, evaluators that use a BSC-format weight financial and non-financial performance differences more heavily than evaluators using an unformatted scorecard. Managers who are involved in selecting strategic initiatives perceive those initiatives afterwards as having been more successful than managers who are not involved in the initiative-selection process. Simply framing the scorecard as a causal chain is not sufficient to mitigate these effects, but framing the scorecard as a causal chain *and* involving managers in the selection of scorecard measures, mitigates the effects of manager involvement in initiative selection on initiative performance evaluation (Tayler, 2010).

Two more studies undertaken in a BSC context investigate how incentive payment affects the evaluation of performance by managers. Choi et al. (2012) remark that firms develop strategic performance measurement systems (SPMS) that translate strategy into imperfect performance measures of the true strategic construct. Unfortunately, managers fail to acknowledge that performance measures are imperfect representations of the strategic construct and act as if the performance measures are the constructs of interest (surrogation). Surrogation is increased by incentive compensation. This effect is larger when compensation is based on a single measure of the strategic construct compared to when it is based on multiple measures of a strategic construct (Choi et al., 2012). Ding & Beaulieu (2011) show that participants who were induced to feel good (bad) gave higher (lower) evaluation scores to divisional managers both in a setting with only two performance measures and in a setting with a 16-measure BSC. Financial incentives eliminated the mood congruency bias in the two-performance-measure- condition and in the condition with a simplified BSC with only eight measures, but not in the 16-measure-BSC-condition. Financial incentives thus can reduce the bias if the BSC is not too extensive (Ding & Beaulieu, 2011).

The literature links judgment biases also to personal characteristics of the subordinate and/or the supervisor. Social and situational influences are important in the performance-rating process. Demographic similarity and the supervisor-subordinate relationship significantly influence supervisors' affection for subordinates and influence performance rating indirectly through supervisors' affection (Judge & Ferris, 1993; Wayne & Liden, 1995). Supervisors appraise the performance of subordinates whom they perceive as being dissimilar to themselves significantly lower (Pulakos & Wexley, 1983) e.g. after controlling for performance, racial differences between subordinate and supervisor lead to lower ratings for both black and white subordinates (Elvira & Town, 2001) and both male and female

supervisors rate subordinates of the same sex higher (Varma & Stroh, 2001). Black subordinates get lower subjective performance ratings than whites (Elvira & Town, 2001). Biernat & Sesko (2013) investigates the evaluations of mixed-sex work teams' performance after having performed a masculine task. A mixed-sex team consists of a white pair, a black pair or a mixed-race pair. Women's competence was solely judged lower in a white pair work team. Black women were not affected by gender bias (Biernat & Sesko, 2013).

Finally, this section discusses the literature on biases related to accompanying or competitive information about the subordinate's performance, i.e. the outcome effect, the assimilation/spillover effect and the halo effect. The outcome effect captures that when an objective outcome measure is positive (negative), supervisors tend to evaluate the subordinate positively (negatively), regardless of the actual appropriateness of the decision resulting in that outcome. For instance, Ghosh & Lusch (2000) document how subjective performance evaluations of store managers were negatively influenced by unfavorable objective outcome knowledge. Similarly, Ittner et al. (2003) document how supervisors tend to focus on outcome-oriented financial performance measures when evaluating subordinates. Taking outcomes into account that do not reflect subordinates' performance will affect not only the quality of the subjective performance evaluation but will also incorrectly reward/penalize subordinates (Ghosh & Lusch, 2000). The assimilation or spillover effect is a bias very similar to the outcome effect. In this case, supervisors bias their subjective evaluations of performance on one dimension to be consistent with an objective measure of performance on a separate and unrelated dimension (Bol & Smith, 2011). Likewise, Duarte et al. (1994) document that ratings are consistent with objective performance measures, however only with low-quality relationships in the short run. The halo effect relates to the observation that the supervisor's prior expectations about a subordinate's performance have an effect on later ratings of that performance. When a subordinate's actual performance disappoints a supervisor's expectations about that performance, subsequent ratings will be lower, regardless whether the actual performance is better or worse than expected. The more often supervisors must use objective measures like rating formats or strict procedures, the more likely ratings are to be accurate (Hogan, 1987). A supervisor's general impression formed from prior interactions may impair the supervisor's ability to objectively assess the subordinate's current work (Tan & Jamal, 2001). More specifically, Tan & Jamal (2001) show in an audit-context that average superiors evaluate work done by outstanding subordinates more favorably than work done by average subordinates when they know the identity of the work preparer, but not when the identity of the work preparer is unknown. Outstanding superiors are not affected by the perceived competence of the subordinate preparing the work. Dossett & Greenberg (1981) investigate how employees can steer this halo effect. They examine the effect of who sets the performance goal (self-set, participative or assigned) and performance outcome on employee's performance evaluation. Their findings indicate that a worker who initially suggested a high goal received a significantly higher performance score than a worker who suggested a

low goal, consistent with the halo effect. Besides, as raters have more opportunity to observe ratee behavior (i.e. higher familiarity between supervisor and subordinate), the magnitude of halo error increases (Jacobs & Kozlowski, 1985). Performance ratings are less accurate when rating is delayed (instead of immediate rating) and when only a small amount of information is observed (Heneman & Wexley, 1983).

V. PERCEIVED FAIRNESS

Whether a performance evaluation is perceived as 'fair' depends amongst others on influence activities (i.e. attempts of subordinates to influence the evaluation of the supervisor) (Du et al., 2012), favoritism (i.e. supervisors acting on personal preferences toward subordinates to favor some subordinates over others beyond their true performance) (Du et al., 2012; Prendergast & Topel, 1996), procedural justice (i.e. the perceived fairness of the means and procedures used to determine the subjective performance evaluations) (Dulebohn & Ferris, 1999; McFarlin & Sweeney, 1992) and whether supervisors adjust their subjective performance evaluations when uncontrollable factors affect subordinates results (Bol & Smith, 2011; Ghosh & Lusch, 2000). Perceived fairness is of utmost importance in subjective performance evaluations. Table 4 summarizes the articles dealing with perceived fairness in subjective performance evaluations. The majority of these articles deal with influence activities, favoritism and procedural justice, while some also deal with controllability.

Table 4. Perceived fairness in subjective performance evaluations

Article	Focus/Results	Research Set-up
Influence activities, favoritism and procedural justice		
Burney, Henle & Widener (2009)	The higher organizational justice, the higher employee performance.	Survey (242 persons, in 47 branches of a large American financial services organization
Du, Tang & Young (2012)	Both influence activities and government favoritism affect the evaluation positively.	Archival study (63 state-owned enterprises (SOEs), 2005–2007) and interviews (6 CFOs of SOEs)
Dulebohn & Ferris (1999)	Supervisor-focused influence tactics are associated with positive procedural justice evaluations, but job-focused influence tactics were associated with negative evaluations.	Field study (128 subordinates and 23 supervisors in a food service department)
Hartmann & Slapnicar (2009)	Subordinate's trust in the superior depends on the formality of the performance evaluation procedure.	Survey (160 departmental managers in 11 Slovenian commercial banks)

Article	Focus/Results	Research Set-up
Hartmann & Slapnicar (2012)	The amount of subordinate's voice in the performance evaluation process has a positive effect on justice perceptions.	Survey (178 upper-middle managers in 12 Slovenian commercial banks)
Hartman, Naranjo-gil & Perego (2010)	Initiating structure-leaders and consideration-leaders both enhance evaluation fairness in their own way.	Survey (196 middle-level managers in 11 Dutch organizations)
McFarlin & Sweeney (1992)	Procedural justice is an important predictor of organizational commitment.	Survey (675 employees of an American bank)
Prendergast & Topel (1996)	Favoritism causes firms to use bureaucratic rules in pay decisions and firms place too little weight on supervisor appraisals. Favoritism reduces incentives because of increased risk in evaluations.	Analytical model
Prendergast (1993)	Subordinates have an incentive to conform to what they feel their superiors want to hear.	Analytical model
Wayne & Liden (1995)	Demographic similarity and subordinates' impression management influence performance ratings.	Survey (111 supervisor-subordinate pairs in nonacademic jobs at 2 American universities)
Woods (2012)	Supervisors use downward adjustments to performance evaluations to encourage the departure of certain subordinates.	Field study (272 observations and 66 surveys in an internal audit organization in 2006)
Controllability		
Bol & Smith (2011)	Supervisors adjust their evaluations when an uncontrollable factor decreases the subordinate's objective measure, but they do not adjust the evaluations when the uncontrollable factor increases subordinate's objective measure.	Experiment (216 non-academic supervising employees at a university)
Ghosh & Lusch (2000)	(Un)controllable outcomes (do not) influence performance evaluations, but central management determinants of outcome, which are uncontrollable, influence evaluations.	Archival study in 204 stores of an American retailer
Gibbs, Merchant, Van der Stede & Vargus (2004)	Subjective bonuses are used to provide employees insurance against downside risk in their pay.	Archival study (526 department managers in 250 American car dealerships in 1998–1999) and 1050 surveys in 326 different dealerships
Giraud, Langevin & Mendoza (2008)	For uncontrollable factors external to the company, managers do not prefer the controllability principle, but for internal, uncontrollable factors, managers prefer the principle.	Survey (265 French managers)

Article	Focus/Results	Research Set-up
Govindarajan (1984)	Superiors of BUs with higher environmental uncertainty will use a more subjective performance appraisal and superiors of BUs with lower environmental uncertainty will use a more formula-based performance evaluation.	Interviews (managers of business units within 8 Fortune 500 firms and 58 surveys)
Höppe & Moers (2011)	Discretionary bonuses are used for risk-reduction purposes.	Archival study (1,753 firm-year- observations for 424 American, publicly listed firms, 1998–2002)
Wong-On-Wing, Guo, Li & Yang (2007)	Top managers do not automatically take into account the quality of strategy (uncontrollable factor for divisional managers) in performance evaluation of divisional units using BSC. Divisional managers automatically consider the quality of strategy without being asked to do so.	Experiment (68 MBA students)

Subjectivity leads to favoritism when evaluators act on personal preferences toward subordinates to favor some employees over others beyond their true performance (Prendergast & Topel, 1996). When other employees discover the favoritism, it leads to a decrease in procedural justice, and associated with that, it results in a decrease in employees' motivation and organizational commitment (McFarlin & Sweeney, 1992; Prendergast & Topel, 1996).

Woods (2012) even provides evidence of how supervisors appear to use downward performance adjustments in order to encourage the departure of certain subordinates. As such, subordinates have an incentive to conform to what they feel their superiors want to hear (Prendergast, 1993). In order to constrain favoritism and to induce subordinates to report information honestly, analytic results show that firms will deemphasize incentive pay for subordinates, increase the use of bureaucratic rules in pay decision and place too little weight on supervisor appraisals, giving too much weight to noncorruptible, objective measures such as seniority (Prendergast, 1993; Prendergast & Topel, 1996). Employees' perceptions of justice in an organizational context increase with the degree to which employees think that the strategic performance measurement system (SPMS) reflects a strategic causal model, the degree to which the SPMS is technically valid[6] (Burney et al., 2009) and the amount of subordinate's voice in the performance evaluation process (Hartmann & Slapnicar, 2012). Subordinate's trust in the superior depends on the formality of the performance evaluation procedure (i.e. procedural justice, represented by explicit targets, clear metrics and clear

[6] A technically valid SPMS provides employees with performance measures information that is accurate, accessible, understandable, reliable and timely. Employees have access to the performance measures information, understand what it means and how to use it in carrying out their job. If this information is used to define employee's incentive compensation, employees will likely think that their evaluations accurately reflect their effort and their expectations (Burney et al., 2009).

bonus allocation rules) as well. This is because formality increases the perceived quality of feedback and perceptions of procedural justice. Formality matters more for trust formation to those managers that are in functions with less contractible outputs (Hartmann & Slapnicar, 2009). Supervisors that score high on consideration leadership style, i.e. supervisors concerned with the promotion of subordinates' well-being through supportive and pleasant relationships, significantly affect procedural fairness directly. Supervisors high on initiating structure leadership style, i.e. supervisors clearly indicating the roles of their subordinates toward the attainment of organizational goals, by in detail deciding what will be done and how it should be done, are effective in enhancing fairness towards their subordinates by clearly communicating expectations and setting objective standards in performance evaluation procedures (Hartman et al., 2010). Burney et al. (2009) show that firms do not necessarily need to introduce subjectivity into the incentive contracts to improve performance. If firms clearly communicate the characteristics of the SPMS incentive plan such that employees perceive this plan has a high degree of technical validity and it highly reflects the organization's strategic causal model, this will enhance employees' perceptions of justice in an organizational context, which will in turn affect employees' performance positively as well (Burney et al., 2009). Evidence in the context of a Chinese government agency evaluating the performance of different state-owned enterprises reveals that both influence activities and favoritism affect performance evaluation positively (Du et al., 2012). Wayne & Liden (1995) develop and test a theoretical model to understand the effect of subordinate impression management[7] (i.e. influence activities) on supervisor performance ratings. They find that demographic similarity and subordinates' impression management influence performance ratings through supervisors' liking of and perceived similarity to subordinates (Wayne & Liden, 1995). Dulebohn & Ferris (1999) investigate the impact of employees' use of influence tactics on their evaluations of the fairness of the performance evaluation process. They distinguish between two categories of influence tactics: supervisor-focused tactics, tactics used by employees to be better liked by their supervisors such as flattery and doing favors, and job-focused tactics, tactics used to self-promote and appear competent. Supervisor-focused influence tactics are associated with positive employee evaluations of procedural justice, but job-focused influence tactics were associated with negative employee evaluations of procedural justice.

The literature on optimal contracting (see section 2) revealed that fairness considerations of supervisors lead to optimal contracts whereby supervisor's subjective performance evaluations are compressed and above average (compression bias and leniency bias) (MacLeod, 2003; Golman & Bhatia, 2012). Bol (2011) empirically investigates the effect of leniency bias on future performance in a financial service provider. She shows that leniency

[7] Impression management are those behaviors individuals employ to protect their self-images, influence the way they are perceived by significant others, or both. The individuals try to look more appealing/favorable to their superior or to peers. This can be accomplished with smiling, eye contact, touching, verbally agreeing, flattery, favor-doing, opinion conformity with the superior or the peer etc. (Wayne & Liden, 1995).

bias increases future performance. This is explained by the fact that subordinates over-estimate their abilities relative to their supervisors. As such, leniency bias results in an increase in congruence between the rating the employee thinks to deserve and the rating the subordinate actually receives. Lenient ratings are more in line with the expectations of self-over-estimating employees and consequently improve perceived fairness of the incentive system and, in turn, employee motivation (Bol, 2011). Moreover, empirical research in a retailer (Ghosh & Lusch, 2000), publicly-listed companies (Höppe & Moers, 2011) and car dealerships (Gibbs et al., 2004) indicates that supervisors take factors uncontrollable to the subordinate but affecting subordinate's performance into account when evaluating subordinate's performance (Ghosh & Lusch, 2000; Höppe & Moers, 2011; Gibbs et al., 2004). Research in eight Fortune 500 firms reveals that superiors of business units which face higher environmental uncertainty will use a more subjective performance appraisal approach and superiors of business units which face lower environmental uncertainty will use a more formula-based performance evaluation approach (Govindarajan, 1984). In the experimental setting of Bol & Smith (2011), supervisors adjust their subjective performance evaluations when an uncontrollable factor decreases the subordinate's unrelated, objective measure (i.e. they compensate for bad luck), but they do not adjust the evaluations when the uncontrollable factor increases subordinate's objective measure (i.e. they do not punish for good luck). This is consistent with fairness considerations (Bol & Smith, 2011). Supervisors provide employees insurance against downside risk in their pay (filtering out uncontrollable factors due to interdependencies, recalculating incentives when performance targets are too challenging or when department is facing losses) (Gibbs et al., 2004). In their field study, Ghosh & Lusch (2000) document how outcome determinants over which subordinates have control influence their subjective performance evaluations and environmental determinants of outcome over which they have no control do not influence their evaluations. However, inconsistent with the reasoning above, Ghosh & Lusch (2000) find that determinants of outcome decided by central management also influence subordinate's performance evaluations, although they have no influence over those determinants. Wong-On-Wing et al. (2007) find a similar result: they find that supervisors do not take into account the quality of strategy, an uncontrollable factor for their subordinates, when evaluating the subordinates unless they are explicitly required to do so. In contrast, subordinates automatically consider the effect of the quality of strategy on their performance without being prompted to do so. When the uncontrollable factors are internal (interdependencies due to decisions made by colleagues or superiors), managers prefer that their supervisors take these uncontrollable factors into account. Yet, at the same time, when the uncontrollable factors are external to the company, managers do not want their supervisors to adjust for these uncontrollable factors. They fear that the procedure to neutralize the effect of the uncontrollable factors may result in subjectivity in an unfair way (Giraud et al., 2008).

VI. AVENUES FOR FUTURE RESEARCH

Based on our literature review, we can identify several avenues for future research. The structure of this section follows the structure of this paper. We start with possible extensions related to optimal contracting. Afterwards we discuss avenues for future research dealing with discretionary bonus pools and we end with sections discussing research opportunities related to 'biases and debiasing' and 'perceived fairness' respectively.

The current state of the literature indicates a number of research opportunities in optimal contracting with subjective performance evaluations. First, older studies such as Gibbs et al. (2004) or Govindarajan (1984) discuss subjectivity in a very general way. These studies just state 'the use of subjectivity' in a broad, general sense without distinguishing between the different forms of subjectivity such as subjective weightings of objective performance measures, the use of subjective performance measures, a subjective performance evaluation by the supervisor or the possibility to take into account non-prespecified factors in the performance evaluation ex post. This research area would benefit from acknowledging that many different types of subjectivity exist. Researchers can extend the classification of different types of subjectivity and investigate the use of more specific types of subjectivity to broaden our understanding of the use, costs, benefits and consequences of different types of subjectivity (Bol, 2008; Du et al., 2012; Höppe & Moers, 2011; Ittner et al., 2003; Ke et al., 1999; Woods, 2012). Second, subjectivity does not occur in isolation; firms make use of both objective and subjective performance measures. Nonetheless, many studies do not take the total incentive contract into account when examining subjectivity. The relationships among the various compensation package elements and the specific situations, in which the reward packages are used, could be investigated (Gibbs et al., 2004). For example, future research could examine whether subjective and objective performance measurement act as complements or substitutes (Bol, 2008). How does rating behavior or the optimal contract of the supervisor differ when both objective and subjective performance measures are used? Research, thus far, has not been able to answer these questions. Third, all agency theoretical models described earlier in this paper are quite simplistic. They can be adapted such that they describe a more realistic setting. E.g., MacLeod (2003) deals with a risk-neutral principal and a risk-averse agent and Budde (2007) only investigates the case with a risk neutral principal and agent. The models of MacLeod (2003) and Budde (2007) would be improved if the principal could be risk-averse as well or when principals and agents could transfer risk from one to the other by making use of their private information (MacLeod, 2003). Also, these models could be extended to the case with multiple subjective evaluators (Baker et al., 1994). Finally, the agency literature has focused extensively on the determinants of optimal weights from the supervisor perspective, but has paid little attention to the implications of those optimal weights on subordinate motivation (Ahn et al., 2010).

Future research can also build on the literature on discretionary bonus pools. According to agency theory, a principal/supervisor designs a contract that maximizes firm value.

However, most firms are multi-layered and the principal/supervisor designing the contract is not the residual claimant. As such, he has little incentive to aim for the optimal, value-maximizing contract. Instead, to some extent personal preferences will introduce subjectivity in the contract design phase (Rajan & Reichelstein, 2009; Baker et al., 1988). In addition, the discretionary bonus pool optimal contracting literature assumes that the principal agrees ex ante to optimally allocating the bonus pool and ex post he has no incentive to do otherwise. However, in real-life the principal might not make the optimal bonus pool allocation due to favoritism, influence activities, collusion among agents or sabotage of the performance of one agent by another agent, which distorts the performance information of that other agent (Baiman & Rajan, 1995; Rajan & Reichelstein, 2006, 2009). Future research could explore the conditions under which these implementation problems arise. For example, future research could investigate what effect the supervisor, being the residual claimant or not, has on the structure of the optimal contract.

In general, researchers could further examine how incentive structures of both principal and agent affect research outcomes (Maas et al., 2012). Research focusing on the circumstances under which we would expect to find positive and negative consequences of performance measurement systems on managerial performance would be beneficial (Franco-Santos et al., 2012). Researchers could also investigate employees' responses (i.e., effort and/or performance) to managers' use of allocation discretion. Employees may be proactive in providing favorable noncontractible information to their superiors, and such behavior may vary across different levels of discretion and/or given their perceptions of managers' allocation processes (Bailey et al., 2011). Additionally, present research focuses on the performance of discretionary bonus pools in a single period. These studies can be extended to a setting with repeated interaction over multiple periods (Rajan & Reichelstein, 2006). In this context supervisors' reputation of trustworthiness might become relevant. Next, we discuss future research opportunities related to the 'biasing & debiasing' stream. In general, there is a need for further research into the role of social and contextual factors in the performance evaluation process (Duarte et al., 1994; Dulebohn & Ferris, 1999; Franco-Santos et al., 2012). The question whether superiors' evaluative behavior could be explained in terms of the context remains unanswered (Hartman et al., 2010). Judge & Ferris (1993) consider the effects of several key aspects of social contexts, such as supervisor-subordinate demographic similarity, supervisor-subordinate work relationship, supervisor's span of control, supervisor's experience, supervisor's affection for the subordinate and supervisor's opportunity to observe subordinate's job performance, on the performance-ratings process. Future research could expand the variables studied and provide a deeper assessment of the causal relationships among those variables, and thus a more informed understanding of the performance-rating process (Judge & Ferris, 1993). Especially research integrating a full range of social context variables with the cognitive processes of supervisors in observing, storing, and recalling data about subordinates is lacking. How do situational variables such as organizational level, work group size, technology and task interdependence influence the processing of information

(Wayne & Liden, 1995)? Future research could also take a look at how different levels of the time delay of the rating and/or of the amount of information provided, affect performance rating accuracy (Heneman & Wexley, 1983). Research could investigate whether a particular bias is influenced by person-specific characteristics (Kane et al., 1995; Moers, 2005; Maas et al., 2012; Pulakos & Wexley, 1983) or social context factors (Duarte et al., 1994; Wayne & Liden, 1995), or both. Both internal factors like cognitive consistency, ego enhancement, commitment, and external ones like organization systems, rewards, and social pressures may explain judgment biases (Hogan, 1987). Given equal circumstances, differences in the tendency to rate could reflect personality or information-processing differences among supervisors. Additional research may reveal that personality factors such as neuroticism, extraversion, openness, agreeableness, and conscientiousness are potential predictors of some rating bias (Kane et al., 1995). In addition, researchers should more closely examine how supervisors' affection for a subordinate, dyadic quality, expectations and history of working together influence the performance appraisal process (Bol & Smith, 2011; Duarte et al., 1994; Hogan, 1987; Judge & Ferris, 1993; Maas et al., 2012). Although several authors have stressed the need to examine gender as an important attribute of social context, only a few field studies have modeled and tested the effects of gender and race composition and supervisor-subordinate similarity on dyadic quality and performance appraisal (Duarte et al., 1994; Elvira & Town, 2001; Varma & Stroh, 2001). Additional work is needed to more fully understand how being a member of a social category produces differential performance evaluation outcomes (Biernat & Sesko, 2013). Minorities are more likely to have supervisors of a different race. So if, differences exist in the performance-evaluation process, ratings may disproportionately disadvantage minority employee outcomes (Elvira & Town, 2001). Until now, little attention has been paid to how a judgment bias can be controlled or reduced as well (Kane et al., 1995; Wong-On-Wing, et al., 2007). The following debiasing suggestions are left for future research. One might examine whether an education in debiasing strategies during training programs for supervisors reduces judgment biases. Such an education would make the supervisors alert for this potential problem in their evaluations of subordinates and may help to overcome, at least to some extent, their biased judgments (Ghosh & Lusch, 2000; Tan & Jamal, 2001; Tayler, 2010). Also the subjective performance measurement literature in the context of the BSC might be extended. Future research might focus on judgment biases when evaluating based on the BSC: how can employees gain a better appreciation for the measures if they are involved in the selection of those measures and the design of the units' scorecard. This might increase the reliance on all BSC measures, including the unique measures and as such the common measure bias might reduce (Lipe & Salterio, 2000; Libby et al., 2004). Future research could also investigate how different presentation formats and features, such as graphs or traffic lights, facilitate the processing of performance information (Cardinaels & van Veen-Dirks, 2010, 2010). Researchers could examine whether the effects of framing the scorecard as a causal chain are stronger when managers are provided with additional causal- chain-related data (e.g., correlations), or when managers are given

additional training on using the causal chain (Tayler, 2010). In addition, if supervisors have to justify their evaluation judgments, this accountability could potentially moderate judgment biases (Bol & Smith, 2011). Another fruitful avenue is to look at the consequences of biased performance ratings on subordinate performance (Bol, 2011; Varma & Stroh, 2001). Future empirical research could, for example, examine whether supervisors rate leniently because they expect that leniency bias positively affects subordinate performance or whether they do so to avoid rating costs (Bol, 2011). Studies that focus on common measure bias can explore whether unique non-financial measures are more easily ignored than unique financial measures in a BSC- format. Evaluators surely tend to focus more strongly on financial measures when measures are organized in a BSC-format (Cardinaels & van Veen-Dirks, 2010). If managers tend to ignore non-financial, unique measures in a BSC-format, these performance measures should not be included in the BSC-framework. As subjective performance measures are often non- financial and unique, companies including these subjective performance measures in the BSC may falsely assume managers take those subjective measures into account.

Finally, we discuss future research opportunities related to perceived fairness. Also in this research area there is a need for further research into the role of social, contextual and person-specific factors (Duarte et al., 1994; Dulebohn & Ferris, 1999; Franco-Santos et al., 2012). Researchers should determine which personal and organizational factors impact procedural justice and how procedural justice, in turn, affects organizational outcomes (McFarlin & Sweeney; 1992). Researchers could analyze the effect of different combinations of performance evaluation system design and use on trust and justice perceptions (Hartmann & Slapnicar, 2009, 2012). Does a superior's reputation of trustworthiness have an impact on employee performance in a subjective performance evaluation setting and does this impact differ depending on the specific subjective performance evaluation system design? Or researchers could explore controllability in more detail. Do evaluators respond differently depending on the type of uncontrollable factor (external or internal factor, economic or competitive factor, natural catastrophe, …) (Giraud et al., 2008)? Do subordinates in a subjective performance evaluation setting respond differently to uncontrollable factors if their superior has a reputation of trustworthiness? Furthermore, future research is needed to define which factors cause favorable or unfavorable supervisor reactions towards influence activities (Wayne & Liden, 1995). Future research could also examine the effects of employees' use of influence tactics on their justice evaluations of the performance evaluation system (Dulebohn & Ferris, 1999).

VII. CONCLUSION

In this article we review academic research on subjective performance measurement in high-impact journals. The final selection consists of 67 articles published in 20 high-impact journals during the period 1977 to 2013. We classified this final selection of 67 articles in four

research streams: 'optimal contracting', 'discretionary bonus pools', 'judgment biases and debiasing', and 'perceived fairness' and discussed each stream in detail. In the section on optimal contracting benefits and costs related to subjectivity in performance contracts are discussed. Subjectivity can tackle perceived weaknesses in quantitative formulaic bonuses such as incompleteness, short-term focus, susceptibility to manipulation, incentive distortions (congruity issues), risk concerns, environmental uncertainty, uncontrollable factors, moral hazard, asymmetric information or agents "gaming" or manipulating the performance evaluation system. However, subjectivity in performance contracts may cause conflicts and disputes between subordinate and supervisor and judgment biases and evaluation uncertainty may harm employee motivation and productivity. The section on discretionary bonus pools makes clear that this specific type of bonus pools (the magnitude of the bonus pool is based on an explicit formula agreed-upon ex ante and the allocation of the bonus amongst the subordinates is based on supervisor's discretion) can face many of the downsides related to subjectivity in performance contracts, while incorporating the benefits related to subjectivity. The section on judgment biases and debiasing discusses that the introduction of subjectivity in performance measurement introduces different judgment biases by supervisors such as compression bias, centrality bias, biases related to the balanced scorecard, biases related to personal characteristics or biases related to accompanying or competitive information. These judgment biases have an effect on subordinate performance and the perceived fairness of the subjective performance evaluation. The section on perceived fairness shows that perceived fairness is of utmost importance in subjective performance evaluation and that influence activities, favoritism, procedural justice and the (non)-existence of adjustments for uncontrollable factors affect perceived fairness. This article ends with a section on avenues for future research, in which we discuss that more research is needed to uncover the use, benefits and costs related to specific types of subjectivity. Furthermore, we stress that subjectivity should not be studied in isolation, but rather as a part of the whole performance measurement system taking into account the interaction between objective and subjective parts of a compensation contract. Also more research on the role of social, contextual and person-specific factors in the performance evaluation process is needed.

REFERENCES

Ahn, T. S., Hwang, I., & Kim, M.-I. 2010. The Impact of Performance Measure Discriminability on Ratee Incentives. Accounting Review, 85(2): 389–417.

Bailey, W.J., Hecht, G., & Towry, K.L. 2011. Dividing the Pie: The Influence of Managerial Discretion Extent on Bonus Pool Allocation. Contemporary Accounting Research, 28(5): 1562-+.

Baiman, S., & Rajan, M.V. 1995. The Information Advantages of Discretionary Bonus Schemes. Accounting Review, 70(4): 557–579.

Baker, G., Gibbons, R., & Murphy, K.J. 1994. Subjective Performance-measures in Optimal Incentive Contracts. Quarterly Journal of Economics, 109(4): 1125–1156.

Baker, G.P., Jensen, M.C., & Murphy, K.J. 1988. Compensation and Incentives – Practice vs Theory. Journal of Finance, 43(3): 593–616.

Banker, R.D., Chang, H.S., & Pizzini, M.J. 2004. The balanced scorecard: Judgmental effects of performance measures linked to strategy. Accounting Review, 79(1): 1–23.

Biernat, M., & Sesko, A.K. 2013. Evaluating the contributions of members of mixed-sex work teams: Race and gender matter. Journal of Experimental Social Psychology, 49(3): 471–476.

Bol, J.C. 2008. Subjectivity in Compensation Contracting. Journal of Accounting Literature, 27: 1–24.

Bol, J.C. 2011. The Determinants and Performance Effects of Managers' Performance Evaluation Biases. Accounting Review, 86(5): 1549–1575.

Bol, J.C., & Smith, S.D. 2011. Spillover Effects in Subjective Performance Evaluation: Bias and the Asymmetric Influence of Controllability. Accounting Review, 86(4): 1213–1230.

Bommer, W.H., Johnson, J.L., Rich, G.A., Podsakoff, P.M., & Mackenzie, S.B. 1995. On the Interchangeability Of Objective and Subjective Measures of Employee Performance – A Metaanalysis. Personnel Psychology, 48(3): 587–605.

Budde, J. 2007. Performance measure congruity and the balanced scorecard. Journal of Accounting Research, 45(3): 515–539.

Burney, L.L., Henle, C.A., & Widener, S.K. 2009. A path model examining the relations among strategic performance measurement system characteristics, organizational justice, and extra- and in-role performance. Accounting Organizations and Society, 34(3–4): 305–321.

Cardinaels, E., & van Veen-Dirks, P.M.G. 2010. Financial versus non-financial information: The impact of information organization and presentation in a Balanced Scorecard. Accounting Organizations and Society, 35(6): 565–578.

Choi, J., Hecht, G.W., & Tayler, W.B. 2012. Lost in Translation: The Effects of Incentive Compensation on Strategy Surrogation. Accounting Review, 87(4): 1135–1163.

Cronqvist, H., & Fahlenbrach, R. 2013. CEO contract design: How do strong principals do it? Journal of Financial Economics, 108(3): 659–674.

Ding, S., & Beaulieu, P. 2011. The Role of Financial Incentives in Balanced Scorecard-Based Performance Evaluations: Correcting Mood Congruency Biases. Journal of Accounting Research, 49(5): 1223–1247.

Dossett, D.L., & Greenberg, C.I. 1981. Goal Setting and Performance Evaluation – An Attributional Analysis. Academy of Management Journal, 24(4): 767–779.

Du, F., Tang, G., & Young, S.M. 2012. Influence Activities and Favoritism in Subjective Performance Evaluation: Evidence from Chinese State-Owned Enterprises. Accounting Review, 87(5): 1555–1588.

Duarte, N.T., Goodson, J.R., & Klich, N.R. 1994. Effects of Dyadic Quality and Duration on Performance-appraisal. Academy of Management Journal, 37(3): 499–521.

Dulebohn, J.H., & Ferris, G.R. 1999. The role of influence tactics in perceptions of performance evaluations' fairness. Academy of Management Journal, 42(3): 288–303.

Elvira, M., & Town, R. 2001. The effects of race and worker productivity on performance evaluations. Industrial Relations, 40(4): 571–590.

Fisher, J.G., Maines, L.A., Peffer, S.A., & Sprinkle, G.B. 2005. An experimental investigation of employer discretion in employee performance evaluation and compensation. Accounting Review, 80(2): 563–583.

Franco-Santos, M., Lucianetti, L., & Bourne, M. 2012. Contemporary performance measurement systems: A review of their consequences and a framework for research. Management Accounting Research, 23(2): 79–119.

Ghosh, D., & Lusch, R.F. 2000. Outcome effect, controllability and performance evaluation of managers: some field evidence from multi-outlet businesses. Accounting Organizations and Society, 25(4–5): 411–425.

Gibbs, M., Merchant, K.A., Van der Stede, W.A., & Vargus, M.E. 2004. Determinants and effects of subjectivity in incentives. Accounting Review, 79(2): 409–436.

Giraud, F., Langevin, P., & Mendoza, C. 2008. Justice as a rationale for the controllability principle: A study of managers' opinions. Management Accounting Research, 19(1): 32–44.

Golman, R., & Bhatia, S. 2012. Performance evaluation inflation and compression. Accounting Organizations and Society, 37(8): 534–543.

Govindarajan, V. 1984. Appropriateness of Accounting Data in Performance Evaluation – An Empirical-examination of Environmental Uncertainty as an Intervening Variable. Accounting Organizations and Society, 9(2): 125–135.

Hartmann, F., Naranjo-Gil, D., & Perego, P. 2010. The Effects of Leadership Styles and Use of Performance Measures on Managerial Work-Related Attitudes. European Accounting Review, 19(2): 275–310.

Hartmann, F., & Slapnicar, S. 2009. How formal performance evaluation affects trust between superior and subordinate managers. Accounting Organizations and Society, 34(6–7): 722–737.

Hartmann, F., & Slapnicar, S. 2012. The perceived fairness of performance evaluation: The role of uncertainty. Management Accounting Research, 23(1): 17–33.

Heneman, R.L., & Wexley, K.N. 1983. The Effects of Time-delay in Rating and Amount of Information Observed on Performance Rating Accuracy. Academy of Management Journal, 26(4): 677–686.

Hoeppe, F., & Moers, F. 2011. The Choice of Different Types of Subjectivity in CEO Annual Bonus Contracts. Accounting Review, 86(6): 2023–2046.

Hogan, E.A. 1987. Effects of Prior Expectations on Performance Ratings – A Longitudinal study. Academy of Management Journal, 30(2): 354–368.

Humphreys, K.A., & Trotman, K.T. 2011. The Balanced Scorecard: The Effect of Strategy Information on Performance Evaluation Judgments. Journal of Management Accounting Research, 23(1): 81–98.

Indjejikian, R. J., & Matejka, M. 2012. Accounting Decentralization and Performance Evaluation of Business Unit Managers. Accounting Review, 87(1): 261–290.

Ittner, C.D., Larcker, D.F., & Meyer, M.W. 2003. Subjectivity and the weighting of performance measures: Evidence from a balanced scorecard. Accounting Review, 78(3): 725–758.

Ivancevich, J.M. 1983. Contrast Effects in Performance Evaluation and Reward Practices. Academy of Management Journal, 26(3): 465–476.

Jacobs, R., & Kozlowski, S.W.J. 1985. A Closer Look at Halo Error in Performance Ratings. Academy of Management Journal, 28(1): 201–212.

Judge, T.A., & Ferris, G.R. 1993. Social-context of Performance Evaluation Decisions. Academy of Management Journal, 36(1): 80–105.

Kane, J.S., Bernardin, H.J., Villanova, P., & Peyrefitte, J. 1995. Stability of Rater Leniency – 3 Studies. Academy of Management Journal, 38(4): 1036–1051.

Kaplan, R.S., & Norton, D.P. 1992. The balanced scorecard--measures that drive performance. Harvard business review, 70(1): 71–79.

Ke, B., Petroni, K., & Safieddine, A. 1999. Ownership concentration and sensitivity of executive pay to accounting performance measures: Evidence from publicly and privately- held insurance companies. Journal of Accounting & Economics, 28(2): 185–209.

Keeley, M. 1977. Subjective Performance Evaluation and Person-role Conflict Under Conditions of Uncertainty. Academy of Management Journal, 20(2): 301–314.

Krishnan, R., Luft, J.L., & Shields, M.D. 2005. Effects of accounting-method choices on subjective performance-measure weighting decisions: Experimental evidence on precision and error covariance. Accounting Review, 80(4): 1163–1192.

Levin, J. 2003. Relational incentive contracts. American Economic Review, 93(3): 835–857.

Libby, T., Salterio, S.E., & Webb, A. 2004. The balanced scorecard: The effects of assurance and process accountability on managerial judgment. Accounting Review, 79(4): 1075–1094.

Lipe, M.G., & Salterio, S. 2002. A note on the judgmental effects of the balanced scorecard's information organization. Accounting Organizations and Society, 27(6): 531–540.

Lipe, M.G., & Salterio, S.E. 2000. The balanced scorecard: Judgmental effects of common and unique performance measures. Accounting Review, 75(3): 283–298.

Maas, V.S., van Rinsum, M., & Towry, K.L. 2012. In Search of Informed Discretion: An Experimental Investigation of Fairness and Trust Reciprocity. Accounting Review, 87(2): 617–644.

MacLeod, W.B. 2003. Optimal contracting with subjective evaluation. American Economic Review, 93(1): 216–240.

McFarlin, D.B., & Sweeney, P.D. 1992. Distributive and Procedural Justice as Predictors of Satisfaction with Personal and Organizational Outcomes. Academy of Management Journal, 35(3): 626–637.

Merchant, K.A., Chow, C.W., & Wu, A. 1995. Measurement, Evaluation and Reward of Profit Center Managers – A Cross-cultural Field Study. Accounting Organizations and Society, 20(7–8): 619–638.

Moers, F. 2005. Discretion and bias in performance evaluation: the impact of diversity and subjectivity. Accounting Organizations and Society, 30(1): 67–80.

Morse, A., Nanda, V., & Seru, A. 2011. Are Incentive Contracts Rigged by Powerful CEOs? Journal of Finance, 66(5): 1779–1821.

Prendergast, C. 1993. A Theory of Yes Men. American Economic Review, 83(4): 757–770.

Prendergast, C., & Topel, R. H. 1996. Favoritism in organizations. Journal of Political Economy, 104(5): 958–978.

Prendergast, C., & Topel, R. 1993. Discretion and Bias in Performance Evaluation. European Economic Review, 37(2–3): 355–365.

Pulakos, E.D., & Wexley, K.N. 1983. The Relationship Among Perceptual Similarity, Sex, and Performance Ratings in Manager-subordinate Dyads. Academy of Management Journal, 26(1): 129–139.

Rajan, M.V., & Reichelstein, S. 2006. Subjective performance indicators and discretionary bonus pools. Journal of Accounting Research, 44(3): 585–618.

Rajan, M.V., & Reichelstein, S. 2009. Objective versus Subjective Indicators of Managerial Performance. Accounting Review, 84(1): 209–237.

Tan, H.T., & Jamal, K. 2001. Do auditors objectively evaluate their subordinates' work? Accounting Review, 76(1): 99–110.

Tayler, W.B. 2010. The Balanced Scorecard as a Strategy-Evaluation Tool: The Effects of Implementation Involvement and a Causal-Chain Focus. Accounting Review, 85(3): 1095–1117.

Varma, A., & Stroh, L.K. 2001. The impact of same-sex LMX dyads on performance evaluations. Human Resource Management, 40(4): 309–320.

Wayne, S.J., & Liden, R.C. 1995. Effects of Impression Management on Performance Ratings – A Longitudinal study. Academy of Management Journal, 38(1): 232–260.

Wong-On-Wing, B., Guo, L., Li, W., & Yang, D. 2007. Reducing conflict in balanced scorecard evaluations. Accounting Organizations and Society, 32(4–5): 363–377.

Woods, A. 2012. Subjective adjustments to objective performance measures: The influence of prior performance. Accounting Organizations and Society, 37(6): 403–425.

THE LINK BETWEEN CORPORATE ENVIRONMENTAL PERFORMANCE AND CORPORATE VALUE: A LITERATURE REVIEW

Roel Brouwers, Frederiek Schoubben, Cynthia Van Hulle and Steve Van Uytbergen*

Abstract

There is a long-standing debate regarding the link between corporate environmental performance and financial firm performance. Up to the present, this debate has been an important trigger for empirical research. It is often argued, however, that the large body of research concerning this topic has not led to conclusive findings. Mixed results invite a literature study that can clarify the debate and allows for the drawing of conclusions. We focus on studies that examine the impact of corporate pollution as well as corporate initiatives to reduce pollution, and this both within a regulated and a voluntary framework. The literature review reveals that regulation does not enhance the relationship between environmental and financial performance. Legislative actions by governmental bodies merely help in generating environmental awareness among stakeholders as well as in creating a benchmark against which good and bad environmental performance can be defined. It is the stakeholders, enforced by increasing environmental corporate disclosure, who truly force firms to adopt more sustainable business models.

Keywords: environmental performance; environmental regulation; firm performance; pollution; voluntary environmental programs

JEL codes: Q51, Q58, C25

* Roel Brouwers: Corresponding author, KU Leuven, Faculty of Economics and Business, Department of Finance, Korte Nieuwstraat 33, 2000 Antwerp, Belgium; email: roel.brouwers@kuleuven.be.
Frederiek Schoubben: KU Leuven, Faculty of Economics and Business, Department of Finance, Korte Nieuwstraat 33, 2000 Antwerp, Belgium; email: frederiek.schoubben@kuleuven.be.
Cynthia Van Hulle: KU Leuven, Faculty of Economics and Business, Department of Accountancy, Finance and Insurance, Naamsestraat 69, 3000 Leuven, Belgium; email: cynthia.vanhulle@kuleuven.be.
Steve Van Uytbergen: KU Leuven, Faculty of Economics and Business, Department of Finance, Korte Nieuwstraat 33, 2000 Antwerp, Belgium; email: steve.vanuytbergen@kuleuven.be.

I. INTRODUCTION

This article provides an overview of the academic literature on the effect of corporate pollution on firm performance. We concentrate on prior research that has investigated the link between corporate pollution as well as corporate efforts to reduce pollution on the one hand, and corporate value on the other hand. The question that is raised in this review is whether the relationship between environmental and financial performance is primarily driven by mandatory environmental regulations or by disclosure efforts in the context of voluntary environmental programs.

A heated and long-running debate has been raging on the impact of corporate pollution on corporate value. The discussion has been stimulated by the tightening of environmental regulations and the increase in investors' environmental awareness, especially since events such as the Rio Earth Summit in 1992 and the Kyoto Protocol in 1997. The link between environmental and financial performance has been explored in a large number of studies, although there is still no agreement about this relationship. Early studies argued that good environmental performance imposes extra costs on firms (Walley and Whitehead, 1994; Palmer et al., 1995) while other, more recent research provides evidence to support a positive link between corporate environmental performance and corporate financial performance (e.g., Konar and Cohen, 2001; Guenster et al., 2011). According to Derwall et al. (2005) improved environmental performance can increase corporate efficiency and thus create a competitive advantage.

Studies on the relationship between corporate pollution and firm performance can be broadly divided into mandatory and voluntary approaches. The fundamental distinction between both approaches lies in polluters being subjected to either explicit costs imposed by legislators in the case of a mandatory program or implicit costs imposed by stakeholders in a voluntary context. Under a voluntary approach, a polluting firm will not participate unless its payoff is at least as high as it would be without participation, meaning that the firm must perceive some gain, or at least no net loss, ensuing from participation. With mandatory approaches regulators are able to impose net costs on polluting firms, thereby making them worse off than they would have been in the absence of the policy (Alberini and Segerson, 2002). Environmental disclosure plays a key role in the effectiveness of voluntary programs as part of the costs/benefits from environmental actions stem from companies' environmental reputations towards their stakeholders. Given the clear difference between mandatory (i.e., regulatory) and voluntary (i.e., disclosure) approaches and its possible impact on the link between environmental and firm performance, we shall distinguish between studies in regulated and voluntary settings.

The economic costs of environmental regulations have been widely debated since the US began to limit water pollution in 1972 through the Clean Air and Water Acts. According to Stewart (1993) it is cheaper for firms to operate in countries where environmental regulation

is not enforced since regulation involves fines, financial obligations and administrative or legal action against polluting companies. There is also some evidence suggesting that environmental regulation affects productivity because it forces firms to commit resources to non-productive activities such as environmental auditing, waste treatment and litigation (Gray and Shadbegian, 1995, cited by Ramiah et al., 2013). Also Sarkis and Cordeiro (2001) and Rassier and Earnhart (2010) provide evidence of a negative relationship between environmental regulation and market value. In contrast to these studies, several authors have argued that environmental regulation creates rather than destroys value. Porter and Van der Linde (1995), for example, question conventional wisdom about the effect of environmental regulation on firm performance by stating that well-designed regulation could improve a firm's competitiveness. This study led to the so-called Porter-hypothesis which postulates that environmental regulation may result in a win-win situation in that it reduces pollution while simultaneously increasing profits.

In addition to mandatory regulations voluntary environmental investments have emerged as important instruments of environmental policy and governance across the world. The central purpose of voluntary environmental approaches is to produce positive externalities beyond the demands of environmental regulation. Firms that voluntarily adopt progressive environmental policies gain credibility by signaling their environmental position to external stakeholders who cannot otherwise fully observe participants' environmental performance (Prakash and Potoski, 2012). There are theoretical arguments both in favor and against the profitability of voluntary environmental efforts. These arguments range from a clear rejection of voluntary environmental investments with them being considered a total waste of money (Friedman, 1970) to a belief that environmental investments not only pay for themselves but also produce a profit in most cases, while improving environmental quality at the same time (Porter and Van der Linde, 1995).

This literature review provides quite consistent evidence of a negative relationship between firms' emissions, both within a regulated and a voluntary framework, and financial performance. Additionally, several studies document a positive association between pollution reduction and firm value. Other studies show that voluntary environmental initiatives mitigate the negative effects of pollution on firm value in the light of stricter environmental regulations. This review therefore confirms the stakeholder theory and resource-based-view theory suggesting that firms can improve their financial performance by satisfying stakeholders' demands and implementing a proactive environmental strategy. More importantly, it also reveals that disclosure reinforces the positive (stakeholder) effect of environmental initiatives while regulation merely offers a benchmark against which environmental behavior can be measured and compared across firms. Results demonstrate that investors view environmentally proactive firms as better prepared to cope with (expected) future environmental regulation.

The remainder of the review paper is organized as follows. In section 2 we present the most important theories regarding the impact of environmental performance on financial

performance. Section 3 reports the empirical research on the link between corporate pollution, within a mandatory regulation framework, and firm value. In section 4 we present the empirical research investigating the relationship between voluntary environmental programs and firm value. Section 5 concludes this review and provides suggestions for future work in the area.

II. THEORETICAL FRAMEWORK

Although scholars have considered different theoretical views to explain the relationship between corporate environmental performance and firm performance, to date theories have been inconclusive and empirical evidence has been mixed. In this section we consider the most prominent views.

According to the neoclassical agency theory, the expected costs of a firm's environmental responsibility are likely to outweigh the resulting profits and, hence, a firm's environmental performance is expected to have a negative impact on its profitability (Friedman, 1970). Aupperle et al. (1985) explain this neoclassical rationale by arguing that firms that invest in pollution control will incur costs that outweigh the financial benefits. As a consequence, corporate environmental investments can lead to reduced profits or competitive disadvantage and may therefore result in lower profit expectations by investors. The principal agency theory argument related to environmental performance is that corporate environmental responsibility can introduce an agency problem between a firm's management and its shareholders. Friedman (1970) asserts that engaging in corporate environmental responsibility is symptomatic of an agency problem or a conflict between the interests of managers and shareholders. He argues that managers use corporate environmental responsibility as a means to further their own social, political, or career agendas, at the expense of shareholders. According to this view, resources devoted to environmental responsibility would be spent more wisely on efforts to increase firm efficiency. In short, this agency problem causes a negative relationship between environmental performance and financial performance.

Under the agency view, environmental regulation as well as voluntary environmental disclosure would only exacerbate the negative link between environmental performance and firm performance. As proactive environmental investments are, according to this agency view, not in the interest of shareholder wealth maximization, environmental regulation can only force managers to invest in negative NPV projects. These projects will at best shield the firm from non-compliance fines which never compensate for the wealth loss let alone other additional regulatory costs. Also environmental disclosure would not mitigate the negative impact of environmental investing as it is considered unable under this view to align the environmental agenda with wealth maximization.

The agency perspective has been challenged by Freeman (1984) who, in the context of the stakeholder theory, pointed out that every corporation has relationships with many stakeholders and that these stakeholders both affect and simultaneously are affected by the firm's actions. These stakeholder groups include internal and external constituents. Like shareholders, the other stakeholders may place demands upon the firm. Firms must address these demands or else face negative confrontations with non-shareholder groups, which can lead to diminished shareholder value, through boycotts, lawsuits, protests, etc. From a stakeholder theory perspective, corporate social and environmental performance is assessed in terms of a company meeting the demands of multiple stakeholders. Satisfying stakeholder demands to at least some extent is considered an unavoidable cost of doing business. In particular, stakeholder theory suggests that corporate environmental performance should be positively reflected in a firm's financial performance. This is based on the argument that serving the implicit claims of various stakeholders will enhance a firm's reputation, which will consequently lead to a positive impact on its financial performance.

Under the stakeholder view, voluntary environmental disclosure will reinforce the positive relationship between environmental and firm performance as it increases the exposure of environmental activity towards stakeholders. The role of environmental regulation would be at best ambiguous under this paradigm. Once the stakeholders are convinced of the benefits of environmental responsibility, stakeholder theory predicts that they will enforce conducive behavior upon the company without legislation. If that is the case, government intervention will only create unnecessary regulatory costs (Blacconiere and Patten, 1994). According to this view, stakeholder pressure exerts a significant influence on firms' implementation of environmental practices while governmental pressures are less relevant (Gonzáles-Benito and Gonzáles-Benito, 2006; Wood and Ross, 2006; Rivera-Camino, 2007).

At best, government intervention through environmental regulation can create environmental consciousness in society as a whole and in a later stage provide a benchmark for defining good and bad environmental behavior (Tietenberg, 1990).

The arguments from stakeholder theory can be embedded into the resource-based view of the firm. This perspective, introduced by Wernerfelt (1984) presumes that firms are bundles of heterogeneous capabilities and resources. Barney (2001) maintains that differences in organizational performance are a consequence of the heterogeneity of a firm's resources. Hart (1997) argues that the resource-based view of the firm provides a theory to explain competitive advantage as an outcome of the development of valuable organizational capabilities, such as continuous innovation and stakeholder integration, associated with a proactive integration of environmental issues into strategic management. In brief, this theory implies that environmental responsibility leads to competitive advantages and enhanced firm value. Ruf et al. (2001) state that the stakeholder theory can be complemented by the resource-based view of the firm. From a resource-based view perspective, firms can meet stakeholder demands as a strategic investment, requiring commitments beyond the minimum that is necessary to satisfy stakeholders. By strategically investing in stakeholders' demands,

firms gain a competitive advantage through the development of additional, complementary skills (Russo and Fouts, 1997). Russo and Fouts (1997) give an example where a firm has two choices to satisfy the stakeholder requirements for mitigating pollution. The firm could invest in end-of-pipe filtering equipment or the firm could change its production process to reduce pollution. Installing the filtering equipment will satisfy the stakeholders' demands. However, the resource-based view of the firm states that a firm that strategically invests in stakeholders' demands by changing its production process, may enjoy a sustainable competitive advantage over a firm that only installs a filtering equipment.

Under the resource-based view, government intervention through environmental regulation would hardly be relevant as the competitive advantage of proactive environmental activity is by definition firm-specific and cannot create value on an aggregate level. The difference in the disclosure effect between plain stakeholder theory and the resource-based view remains an empirical question as it is not a priori clear whether every type of environmental activity would benefit from increased transparency.

Next to the agency view, implying a negative link, and the stakeholder/resource-based view, suggesting a positive link, McWilliams and Siegel (2001) propose a neutral relationship between environmental and firm performance. They argue that a firm's optimal level of investment in social environmental responsibility can be assessed in an identical way as any other investment by considering demand and supply sides. According to McWilliams and Siegel (2001) firms that do not invest in corporate environmental responsibility will offer their products at lower prices while those firms which incur environmental costs will be able to sell their products at higher prices. Therefore, the relationship between corporate environmental performance and financial performance is expected to be neutral.

The previous discussion shows that the literature is dominated by two opposing views. The negative link between environmental and firm performance is explained by the agency theory whereas the positive link is supported by the stakeholder and resource-based view perspectives. The agency theory implies that environmental responsibility is a misuse of corporate resources that would be better spent on value-added internal projects or returned to shareholders. It also suggests that managers use corporate social responsibility to advance their careers or other personal agendas. Stakeholder theory presents a more positive view on environmental responsibility. This theory asserts that managers must satisfy a variety of stakeholders (e.g., workers, customers, suppliers, local community organizations) who can influence firm outcomes. According to this view, it is not sufficient for managers to focus exclusively on the needs of shareholders. The resource-based view complements stakeholder theory and states that firms should move from mere compliance to active support of stakeholders' environmental requirements to gain a competitive advantage over their competitors. Overall, stakeholder and resource-based view theory suggest that "it pays to be green". In addition to the negative agency and the positive stakeholder/resource-based view there is the neutral approach proposed by McWilliams and Siegel (2001) who argue that firms simply supply a certain demanded level of environmental performance to maximize their profits.

III. ENFORCING THE ENVIRONMENT: MANDATORY ENVIRONMENTAL LEGISLATION

This section reviews research on the effects of environmental regulation on firm performance as well as studies using environmentally regulated emissions as a proxy for environmental performance.

Historically, policymakers have relied on regulatory restrictions on polluting behavior to guarantee adequate protection of environmental quality. The key theoretical argument for environmental regulation is that pollution is a classic example of an externality, an unintended result of market decisions, which affects individuals other than the decision-makers. Because firm-level decisions do not take into account full social costs, pollutant emissions tend to be higher than socially efficient levels. As environmental quality is thus naturally underprovided for by competitive markets, a potential role arises for environmental regulation (Revesz and Stavins, 2003).

We use a geographical dimension to structure the studies within a mandatory context and discuss them in chronological order starting with the research on US mandatory programs, followed by an overview of the literature on the relationship between carbon performance within the European Union Emission Trading Scheme and firm value. Next we consider some studies on the Australian emission reporting scheme. We end this section by examining the effect of Japanese and Chinese environmental legislation on firm value.

A. US LEGISLATION: FROM THE CLEAN AIR WATER ACT TO THE SO_2 EMISSION TRADING SCHEME

Since 1969 several legislative actions in the United States have been aimed at reducing industrial pollution.[1] Simultaneously the US Securities and Exchange Commission has been engaged in developing pollution disclosure requirements to ensure sufficient disclosure of pollution information (Jaggi & Freedman, 1992).

The impact of US environmental regulation on firm performance has been the subject of many studies. Several studies have analyzed the impact of pollution disclosures within the context of early US environmental regulations. Shane and Spicer (1983), for instance, investigate whether stock price movements are associated with the release of externally produced information about companies' performances in the pollution-control area. Specifically, the study examines stock price movements associated with the disclosure of eight studies conducted by the Council on Economic Priorities (CEP)[2] with regard to firms' environmental performances. The results show that pollution disclosures were associated

1 See Carriker (1996) for an overview.

2 The CEP is a research organisation, founded in the USA, that analyses and reports on the social and environmental records of companies.

with market reactions. However, the study of Shane and Spicer (1983) merely analyzes whether the pollution information has *an* impact on market value and not whether there was a positive or negative reaction. A study investigating the signed impact of pollution information on firm performance is Jaggi and Freedman (1992). The authors develop a pollution index based on pollution levels for 13 pulp and paper firms and link the index to economic performance indicators by using the Pearson Correlation test. The results show a negative association between environmental and economic performance and the authors suggest that in the short run a firm's profitability is negatively affected by pollution abatement activities involving high expenditures. This finding provides some support for the neoclassical view that abatement activities are a misuse of firm resources that would be more wisely spent on efforts to increase firm efficiency.

Blacconiere and Northcut (1997) use the event study methodology to examine market reactions for 72 chemical firms to announcements of legislative events leading to the Superfund Amendments and Reauthorization Act (SARA). This act increased direct taxes affecting chemical firms, and expanded regulatory disclosure requirements for firms that release hazardous materials into the environment. The study shows that, due to increased regulatory costs, chemical firms' stock prices had an overall negative reaction to announcements of specific legislative actions (e.g., votes by Congress) leading to the SARA. Furthermore, Blacconiere and Northcut (1997) examine firm-specific environmental information and find that firms with more extensive exposure to regulatory costs experienced a more negative market reaction. These results are in line with the agency perspective that environmental compliance is just an extra cost with no added value.

Since 1987, all US manufacturing facilities with at least 10 employees and producing more than certain benchmarks of each of the 320 listed chemicals, are required to annually report an inventory of toxic releases to the Environmental Protection Agency (EPA). Information about these releases is then publicly disseminated through the Toxics Release Inventory (TRI). Such a requirement informs the public and allows individuals to minimize or avert exposure to toxic substances (Saha and Mohr, 2013). Since TRI is a publicly available database, it is an important metric for stakeholders to measure a company's waste generation and pollution reduction activities across a wide range of industries. Patten (2002) argues that, in support of the use of the TRI data as a proxy for environmental performance, it has the distinct advantage of being based on the same measure for all reporting firms and of covering a large diversified set of firms. These factors have led many scholars to rely on TRI data as the environmental performance indicator of choice.

Hamilton (1995) and Khanna et al. (1998) use event studies to investigate the announcement effect of the yearly TRI data releases. Hamilton (1995) assesses the market reaction to the first release of the TRI data in 1989. For a sample of 436 firms with TRI data, he documents that the average abnormal return on the day the emission information was made public, was equal to -0.284% and statistically significant. In addition, Hamilton (1995) shows that firms reporting TRI information lost on average $ 4.1 million in equity value on

the first day the data were released. This is remarkable as the TRI data releases do not involve any explicit costs. Khanna et al. (1998) study the influence of repeated public disclosures of TRI data over the period 1989–1994. Using a sample of 91 firms, they find significant abnormal returns for the day following the release of TRI data for the years 1991–1994. Contrary to Hamilton (1995) they do not find significant abnormal returns following the first release of TRI data. The difference between Hamilton's (1995) results and the Khanna et al. (1998) study could be attributable to differences with respect to the samples of firms being analyzed. Hamilton (1995) studies a sample of 436 firms in the manufacturing sector of which 12% were in the chemical industry. Khanna et al. (1998) on the other hand use a sample of chemical firms that are generally known to be large polluters relative to firms in other industries. Khanna et al. (1998) argue that a first disclosure of environmental information may not generate significant reactions among investors as chemical firms are known to be large polluters. Repeated disclosures of environmental information however, do lead to statistically significant abnormal returns because repeated provision of environmental information allows investors to benchmark a firm's pollution level and make comparisons of performance over time as well as across firms. Khanna et al. (1998) document that firms whose releases increased relative to the previous year or whose pollution levels rose relative to other firms were confronted with significant negative returns. The firms with decreased pollution levels relative to the previous year or relative to other firms reported insignificant returns.

These two studies examining the impact of TRI disclosures on stock prices may struggle with construct validity issues, however. It seems possible that same-day stock price movements probably reflect contemporaneously reported pollution rankings. These rankings are strongly affected by company size and industry choice and thus the stock market effect could be the result of temporary bad press rather than a real change in perception of a firm's long-term value effect of environmental performance. Perhaps for this reason, these TRI event studies have showed inconsistent evidence in a 5-day window following the TRI data release (King and Lenox, 2001).

Another way to explore the link between environmental and financial performance, as opposed to event studies, is to use standard regression techniques to evaluate the effect of changes in pollution on changes in firm performance. This is in essence the methodology used in a study by Hart and Ahuja (1996). Based on a sample of 137 firms from the S&P 500, they show that changes in pollution over the years 1989–1992 were associated with changes in firm performance as proxied by return on sales, return on assets and return on equity. Their proxy for environmental performance, however, conflates reduction of emissions and divestiture of polluting operations, making it difficult to assess the true value of the effect. This raises the question whether it pays to be green or whether it pays to operate in clean industries. To help distinguish the effect of pollution reduction from other underlying factors, Cohen et al. (1997) construct two portfolios of low polluting and high polluting firms in their respective industries, based on the industry categories used to classify companies in

the S&P 500 index, and compare accounting returns, measured using return on assets and return on equity, and stock market returns between these portfolios. The authors conclude that investors who choose the environmental leaders in an industry-balanced portfolio do just as well as (or better than) investors choosing the environmental laggards in each industry. King and Lenox (2001) differentiate between pollution performance and divestiture of operations in dirtier industries by splitting environmental performance into two constructs: relative performance within one's industries and the average performance of the industries in which one chooses to operate. For an unbalanced sample of 652 firms constituting 4483 firm-year observations for the years 1987 to 1996, they find evidence of a positive association between pollution reduction and financial gains, as proxied by Tobin's Q. They fail, however, to derive the direction of causality.

A second study that links environmental performance to Tobin's Q is Konar and Cohen (2001) who relate the market value of 321 S&P 500 firms to environmental performance, as proxied by TRI emissions and environmental lawsuits. After controlling for variables traditionally thought to explain firm-level financial performance, they find that poor environmental performance is associated with lower Tobin's Q values. Clarkson and Li (2004) follow a different approach. Instead of linking emission levels to firm performance, they examine the market valuation of environmental capital expenditures related to pollution abatement using a modified version of the Ohlson (1995) valuation model.[3] Based on a pooled sample of 256 firm-year observations from 29 pulp and paper firms, their valuation evidence indicates that there are incremental economic benefits associated with environmental capital expenditures by high-polluting firms but not low-polluting firms. The negative association between TRI emissions and firm value on the one hand and the positive link between corporate environmental efforts and firm performance on the other hand suggest that good environmental performance increases the market value of the firm. From a stakeholders and resource-based view perspective the increased market value can be explained by the enhanced corporate reputation which is built on the perceptions of its relevant stakeholders. A strong corporate reputation is both an intangible asset and a source of strategic advantage enhancing a corporation's long term ability to create value (Caves and Porter, 1977).

Studies that analyze the impact of US environmental regulation on firm performance have predominantly been undertaken using either valuation models or event studies (e.g., Hamilton (1995); Clarkson and Li, 2004).

Cordeiro and Sarkis (1997) use a different approach and explore the relationship between TRI data and security analyst earnings forecasts as an alternative proxy for firm performance. For a sample of 523 US firms in 1992, they document a significant, negative relationship

[3] The Ohlson (1995) valuation model is used to estimate abnormal returns and valuation equations. This model is based on the assumption that the market value of a firm's equity can be expressed in terms of a set of 'information variables'. These variables normally include the firm's accounting earnings, the book value of its equity and a variable which captures all other value relevant information not reflected in the firm's accounting records (Ataulah et al. (2006)).

between the level of TRI emissions and industry analyst 1- and 5-year earnings-per-share performance forecasts. Alternatively, Connors and Silva-Gao (2008) explore the "does it pay to be green" question by focusing on the effects of pollution performance on firm-specific risk. The authors examine whether improved environmental performance, measured as reduced TRI emissions, reduces the cost of equity capital. The results indicate that companies with high TRI emissions have a significantly higher cost of equity capital than those with lower emissions when controlling for beta, leverage, information risk, firm size and growth. These results provide evidence that environmental performance, a non-financial performance measure that is receiving growing public exposure, is reflected in the cost of capital. The negative relationship between environmental performance and cost of equity capital is confirmed by Clarkson et al. (2011a) who examine the relevance of environmental disclosures. The authors conclude that TRI emissions are positively associated with the cost of equity capital but that there is no association between voluntary environmental disclosure, which will be discussed in section 4 of this review, and the cost of equity capital. The negative link between environmental performance and the cost of capital strengthens the stakeholder argument as improved environmental performance enhances a firm's reputation which may result in reduced risk and consequently a reduced cost of capital (Miles and Covin, 2000).

The use of TRI data to proxy for environmental performance has its limitations, however. First, toxic emissions represent only one aspect of environmental performance since these emissions do not give any information about the pollution with non-toxic substances such as carbon dioxide emissions (Ziegler et al., 2007). Second, TRI is an aggregate measure of 320 chemicals in which the chemicals are not weighted according to relative risk or physical damage (Ilinitch et al., 1998). Third, as production processes and pollution propensity differ across industries, TRI data from different industries are not easily comparable. Fourth, TRI data rely on self-reported emissions. Finally, the EPA reports TRI data with a two-year lag and the data sometimes represent estimated instead of actual emissions. These limitations could bias parameter estimates when conducting multivariate regressions. However, as the advantages outweigh the limitations, the TRI has gained widespread acceptance within the literature and is most widely used as an empirical proxy for environmental performance (Toffel & Marshall, 2004).

Another proxy that has received some attention within the literature involves the sulphur dioxide (SO2) emissions (Hughes, 2000; Johnston et al., 2008). The SO2 emissions are subject to the first emission cap-and-trade system introduced in the US with the passage of Title IV of the 1990 Clean Air Act (CAA).[4] A cap-and-trade system places a cap, or ceiling, on the aggregate emissions of a group of regulated sources by creating a limited number of tradable emissions allowances for a given period and requiring firms to surrender a quantity of allowances equal to their emissions during that period. The system imposes no particular limits on emissions from any given firm or source. A firm may emit as much as it chooses, as

[4] Title IV of the 1990 Clean Air Act is available on www.epa.gov/air/caa/title4.html.

long as it obtains sufficient allowances to do so. The government may initially distribute the allowances for free or sell them at auction. In either case, the need to surrender valuable allowances to cover any emissions and the opportunity to trade those allowances establishes a price on emissions. In turn, this price provides firms with an incentive to reduce their emissions that influences all of their production and investment decisions (Stavins, 2007).

Hughes (2000) uses SO2 emissions to examine the value relevance of future environmental liabilities of electric utility companies. Using a balance sheet-based valuation model, this study shows that, on average, exposure to unbooked environmental liabilities decreased the mean 1990 share price of electric utilities by 16.3 percent. Hughes (2000) concludes that nonfinancial measures such as toxic emissions are considerably informative to stakeholders and that these measures are impounded into the stock price. Johnston et al. (2008) extend Hughes (2000) by examining allowances that are held in excess of current emission levels. Furthermore they investigate stock price reactions to events involving purchases of emission allowances during auctions sponsored by the United States Environmental Protection Agency (USEPA). The authors conclude that the capital market assigns a positive price to a firm's bank of SO2 emission allowances consistent with the argument that emission allowances have an asset value component that is assigned a positive price by the market.

To summarize this section, it can be stated that the empirical evidence on the value implications of environmental performance within the context of US environmental regulation is quite consistent and convincing.

Most studies find pollution to be negatively related to firm financial performance. Furthermore, pollution reduction seems positively correlated to firm value, providing evidence for the stakeholder theory and the resource-based view. Studies showing a negative link between pollution reduction and firm value are scarce and dated, implying that the stakeholder and resource-based view arguments have been strengthened over time due to growing environmental awareness amongst stakeholders. From this viewpoint environmental regulation becomes less important as stakeholders place pressure on firms to adopt proactive environmental practices that improve their environmental performance. Instruments such as the Toxics Release Inventory and the SO2 cap-and-trade system stay however useful as these systems encourage the public's awareness over environmental issues and develop a benchmark for environmental performance.

B. EUROPEAN UNION EMISSION TRADING SCHEME

The European Union (EU) launched an EU-wide emissions trading scheme (EU ETS) for carbon emissions in 2005 which can be considered as the cornerstone of the EU climate policy. As already discussed in the previous section, the practice of emissions trading is not particularly novel as trading of sulfur dioxide began in the United States in the 1990s (Burtraw et al. 2005). Carbon trading however, which refers to the trading of six major

greenhouse gases[5], is more recent. The EU ETS was the first and is to date the biggest international system for trading greenhouse gas emission allowances, covering almost half of EU's greenhouse emissions and operating in 31 countries (European Commission, 2013).[6] The EU ETS has been designed to operate in different phases. Phase 1 ran between 2005 and 2007 and could be regarded as a start-up and test period. Phase 2, which comprised the years 2008 to 2012, coincided with the Kyoto Protocol commitment period and required EU Member States to achieve an 8% emission reduction compared with their 1990 level. Phase 3 has the longest compliance period, from 2013 to 2020. Its target is to reach by 2020 an emissions level of 21% less than the 2005 level (Mnif and Davison, 2012).

Although the literature on various aspects of the EU ETS is growing, only a limited number of studies have examined the link between carbon performance and firm performance. Anger and Oberndorfer (2008) examine the impact of carbon performance, measured as the allocated carbon emissions divided by actual carbon emissions, on competitiveness defined as a firm's ability to sell and approximate this ability by firms' market revenues. Additionally they analyse the impact of the EU ETS on employment. Applying a regression analysis for 419 German ETS firms, this study reports no influence of carbon performance on revenues or employment.

Schmidt and Werner (2012) use an event study to examine the impact of announcements on actual carbon emissions by the European Commission on stock prices. The abnormal returns, representing the market reaction, is then linked to carbon performance variables such as actual carbon emissions and over-allocation, defined as the difference between allocated and actual emissions. Using a sample of listed firms from Austria, Denmark, Germany and the UK, this study finds a significant link between abnormal returns and over-allocation for two announcement events, providing some support for the hypothesis that firms with over-allocation are rewarded by investors.

The value relevance of emission allowances is examined by Clarkson et al. (2014) who use an Ohlson valuation model and base their study on a sample of 843 firm-year observations over the period 2005–2010. Measuring the firm's pollution level as its shortage in emission allowances, they find a negative relation between this measure and market value. The authors also find that the negative association between firm values and carbon emission shortfalls is mitigated for firms with better carbon performance relative to their industry peers and for firms improving their environmental performance. These findings are consistent with the notion that the market not only bases its assessment on the firm's current emissions profile but also on its anticipated future profile. These findings provide evidence for the stakeholders theory.

---

5 These are carbon dioxide (CO_2), methane (CH_4), nitrous oxide (N_2O), hydrofluorcarbons (HFCs), perfluorocarbons(PFCs and sulphur hexafluoride (SF6).

6 The EU ETS operates in the 27 EU countries, the three EEA-EFTA states (Iceland, Liechtenstein and Norway) and Croatia (joined in 2013).

Overall, a limited number of studies have examined the carbon performance, within the EU ETS, providing some evidence that a surplus of emission allowances is positively associated with firm performance. These results are in line with the findings of Johnston et al. (2008), discussed in the previous section. According to Fornaro et al. (2009) companies face increasing pressure from their stakeholders to measure, disclose, monitor and manage carbon emissions. Therefore, the positive link between excess allowances and firm value is not surprising as firms address the demands of various stakeholders by complying with carbon emission regulation. First, regulated firms that do not meet the emission limits could suffer a loss of reputation as these firms are perceived by stakeholders as failing to address climate risks (Hrasky, 2011). Second, there is a value effect as trade in emission allowances gives value to reducing CO_2 emissions and has formed a market with an asset value worth tens of millions of Euros annually (Miclaus et al., 2008).

C. THE AUSTRALIAN ETS

The effect of environmental regulation on Australian firms' performance is an important issue, given the level of commitment Australia has assigned to green policies, particularly with respect to climate change.[7] It is also important as Australia is one of the largest per capita producers of greenhouse emissions (Ramiah et al., 2013). Yongqing et al. (2013) investigate the potential impacts of an emission trading scheme, which was approved by the Australian government in November 2012. Based on a sample of 200 Australian Securities Exchange indexed companies from 2006 to 2010 and using Heckman's (1979) two-step approach to control for the self-selection bias, this study finds that asset values and operating cash flows will be adversely affected by the reduction plan. Specifically, the book value of long-lived assets is found to be negatively associated with listed companies' carbon emission levels. In addition, this study shows that operating cash flows will decrease for emissions-liable companies. Although operating cash flows have been found to be negatively correlated with emission levels, the empirical results document that this influence is not significant. Yongqing et al. (2013) use the carbon emission data disclosed by the Carbon Disclosure Project (CDP), which is an independent not-for-profit organization with the largest available database of corporate climate change information worldwide. The CDP data have been used in several research papers (e.g., Freedman and Jaggi, 2009; Matsumura et al., 2011) and will be discussed in further detail in Chapter Four.

Chapple et al. (2013) use the event study approach to examine the impact of the proposed Australian emission trading scheme on market value and identify five distinct information events argued to impact the probability of a proposed emission trading scheme being enacted and find evidence that the capital market evaluates the impact of the proposed ETS on firm performance. For a sample of 58 firms with available greenhouse gas emissions data, listed on

7 See Bates (2010) for an overview.

the Australian Stock Exchange, the authors find a negative reaction to all four events argued to increase the likelihood, although only one was statistically significant, and a significant positive reaction to the one event argued to decrease the likelihood of an emission trading scheme. Further, when the authors divide their sample into high and low carbon-intensive firms, they detect a stronger market reaction for the most carbon-intensive firms. Chapple et al. (2013) use data from the Australian National Pollutant Inventory (NPI) to proxy for environmental performance. The NPI works in the same manner as the US Toxic Release Inventory and requires all installations that emit above threshold levels to submit annual reports that quantify their emissions of various land, water and air pollutants. This information is disclosed on the NPI website. Thus, the advantages of using TRI to assess environmental performance also apply to NPI. However, the NPI database has an additional advantage in that it assigns a total risk score to each substance and reports the emissions of individual substances for each facility. These risk scores reflect the relative risk that the substance poses as a function of its environmental hazard, human health hazard and likelihood of exposure to the Australian population or environment (Clarkson et al., 2011c). Ramiah et al. (2013) extend the study of Chapple et al. (2013) and investigate the impact of 19 announcements of environmental regulations using a sample of 1770 companies over the period 2005–2011. Additionally, they estimate the change in systematic risk following the announcements. Remarkably, this study shows that the wealth of shareholders in the electricity industry did not change, which indicates that the biggest polluters are not affected by the introduction of green policies. The authors make the plausible assumption that polluters are passing higher costs to consumers. The shareholders of other industries that are not viewed as the biggest polluters experienced value destruction, with no compensation for these industries. These findings conflict somewhat with the results of Yongqing et al. (2013) and Chapple et al. (2013) who detect a more pronounced negative reaction to the proposed emission trading for companies which are considered to be carbon-intensive.

Overall these studies show that the proposed Australian ETS is relevant to investors and is perceived as having a negative impact on the market value of polluting firms. Further research is needed to assess the actual impact of the scheme and the effect of polluting abatement activities on the firm performance of Australian firms.

D. GREEN POLICIES IN JAPAN AND CHINA

Political commitment to mitigate climate change is increasing throughout Asia, a region accounting for nearly half of the world's carbon emission in 2010. Numerous laws, regulations, procedures and initiatives on protection of natural resources and the environment have been enacted by Asian governments (Calvin et al., 2012). Examples include the Chinese Environmental Impact Assessment Law of 2003[8] or the Japanese Basic Anti-Global Warming

[8] See Bao et al. (2004) for a discussion.

Law of 2010.[9] Thus, Asia is an important region to consider in any discussion of the impact of environmental regulation on firm performance.

The effect of environmental regulation on Asian firms' performance however, has only scarcely been studied. In the Japanese context, Nishitani and Kokubu (2011) and Nishitani et al. (2011) use panel data on manufacturing firms listed in 2009 on the First Section of the Tokyo Stock Exchange, which meet the reporting requirements of Japan's mandatory GHG accounting and reporting system. Using data on 641 Japanese manufacturing firms in the period 2006–2008, Nishitani and Kokubu (2011) examine the influence of firms' reductions of greenhouse gas (GHG) emissions on firm value, measured by Tobin's q. In order to further explore this relationship, this study analyzes not only the effect of the reduction of GHG emissions on firm value but also that of the market discipline imposed by the investors in terms of the reduction of GHG emissions. The empirical estimations prove that firms' reduction of GHG emissions enhances their firm value but also that this occurs where the market discipline imposed by the investors is strong, providing evidence for the stakeholders theory. Nishitani et al. (2011) analyze how a firm's reduction of GHG emissions affects its economic performance, based on a theoretical model derived from the Cobb-Douglas production function and the inverse demand function. In line with the results of Nishitani and Kokubu (2011) the results show that the reduction of GHG emissions increases a firm's economic performance.

Given that China has been the world's second-largest carbon emitter for years and has even overtaken the USA as the world's largest emitter since 2007, environmental protection and energy saving have become important issues in the Chinese government's and firm boards' agenda (Tsang and Kolk, 2010). So far, the only study examining Chinese firms' environmental protection efforts on firm performance is Ye et al. (2013). This paper studies the impact of energy saving efforts on the market values of firms by using the announcement date of the carbon emission rights trading scheme (CERTS) in China as an exogenous shock. The results, based on an event study with a sample of all listed firms in China's Shenzhen Stock Exchange (SZESE), indicate that the efforts of firms on environmental protection were positively valued when the emission trading scheme was introduced.

These results suggest that investors pay increasing attention to efforts related to corporate environmental protection and that firms with more investments in energy saving attract more investors and have more potential increments on their firm value within the framework of an emission trading scheme. The positive effect of environmental protection efforts, enhanced by the introduction of the CERTS, on firm value is in line with the stakeholders and resource-based view.

9 See Adachi and Kikuyama (2010).

IV. ENVIRONMENTAL SELF-REGULATION: VOLUNTARY ENVIRONMENTAL PROGRAMS

Traditionally, the predominant approach to addressing the environmental problems emanating from the private sector was through mandatory environmental regulation. This instrument for environmental protection has however been increasingly criticised for not providing satisfactory answers to the complex environmental problems that society now faces (Annandale et al., 2004). A consequence of this criticism has been the growth of a variety of self-regulating voluntary instruments as supplements and potential alternatives to traditional regulatory approaches. One such instrument, the voluntary environmental program (VEP), is designed to provide participants with incentives to improve their environmental performance. By the late 1990s there were over a dozen voluntary environmental programs and thousands of participating firms in the United States, and this number has continued to grow (Videras and Alberini, 2000).

According to the stakeholders theory, discussed in section 2, voluntary environmental programs have a positive impact on firm value as firms receive a variety of benefits in return for their participation. These benefits include the publicity aspect (Videras and Alberini, 2000), the possibility to signal its willingness to satisfy stakeholders' environmental demands (Ruf et al., 2001) and obtaining goodwill and standing with critical stakeholders (Darnall et al., 2009). The resource – based view provides several additional reasons for firms to join VEPs including the ability to signal the proactivity in its environmental management, therefore indicating that they are greener and cleaner than non-participants (Darnall and Carmin, 2005) and to gain a competitive advantage (Videras and Alberini, 2000).

On the other hand, in the context of the agency theory, participation in VEPs as an indicator of corporate social responsibility, is indicative of self serving behaviour on the part of managers and thus reduces shareholder wealth (McWilliams and Siegel, 2006).

Most of the empirical literature has focused on mandatory disclosure programs as discussed in the previous section. Voluntary disclosure has received less empirical attention, perhaps because of severe self-selection problems when analyzing data (Kim and Lyon, 2011). Most of this work is in the accounting literature and aims to explain the extent of attention to environmental matters in corporate annual reports and corporate social responsibility reports (e.g., Patten, 2002). Little is known about the extent to which voluntary disclosures directly affect financial performance. In this section we review this limited research. We start by focusing on studies linking the Carbon Disclosure Project data to market data followed by a discussion of other VEPs and their impact on firm value. We end this section by analyzing three studies that assess the impact of corporate environmental disclosure on market value.

A. CARBON DISCLOSURE PROJECT

The Carbon Disclosure Project (CDP) is a United Kingdom's based not-for-profit-organization, formed in 2000 as a United Nations initiative. Its mission is to gather and disseminate climate change information in an effort to create a unified response against global warming (Carbon Disclosure Project, 2013). To achieve this goal the CDP enlists the support of signatory institutional investors. Each year the CDP sends a questionnaire to the largest global companies requesting climate change information on behalf of these institutional investors. The results are accessible by the investors as they are received by the CDP and publicly released between September and December.

The data disclosed by the CDP offers some distinctive opportunities for voluntary environmental disclosure research as compared to environmental information released in annual reports and sustainability reports. As Stanny and Ely (2008) indicate annual and sustainability reports are very broad in nature. They include information about all major sustainability issues, making inferences about specific environmental issues very difficult. Furthermore, annual reports include financial information and as Cormier et al. (1993) point out, any market reaction could be caused by correlated relevant financial information. The data disclosed by the CDP is information specifically related to the issue of climate change and the reports are publicized independent of annual reports or other financial information. Recently researchers have started to use CDP data to examine the impact of participation in the CDP and the effect of carbon emissions on firm performance. From the stakeholder perspective, participation in the CDP should be rewarded by investors as CDP participants meet the stakeholders' demand for environmental disclosures (Clarkson, 1995).

Kim and Lyon (2011) examine the effect of CDP participation on share prices for a sample of FT global 500 companies by using the event study methodology. They compare the abnormal returns surrounding CDP disclosures during the period 2003–2006 between participants and non-participants. Contrary to expectations, no evidence is found that participation increased shareholder value. However, by making use of Russia's ratification of the Kyoto Protocol on October 22, 2004, which caused the Protocol to go into effect in all the nations that had ratified it, Kim and Lyon (2011) find that companies' CDP participation increased shareholder value. These results imply that Russia's ratification increased the pressure on the USA and other countries that had not yet ratified Kyoto, to take action on climate change. Consequently, firms in such countries saw the probability of environmental regulation rise. Investors apparently viewed CDP participants as better prepared to deal with climate regulations, leading to increased market value (Kim and Lyon, 2011). While Kim and Lyon (2011) are the only researchers to examine the announcement effect of CDP participation, there are several studies that have considered the link between firms' carbon intensity and its market value by using CDP data. The CDP studies are conceptually similar to those reported in section 3 of this literature review with the only difference that the former uses self-reported CDP data as a proxy for carbon intensity while the latter is based on emission data, subject to

environmental regulation. As a consequence these studies should yield similar results, given that the self-selection bias is properly accounted for.

Aggarwal and Dow (2011) study the effect of greenhouse gas emissions on firm value, as proxied by Tobin's Q, for a sample of 621 large firms from the US, Canada and Europe that filed 2008 emission data with the Carbon Disclosure Project. They conclude that carbon intensity is negatively associated with firm value. Furthermore they obtain information about firms' emission mitigation strategies from the CDP questionnaires and document a positive, although insignificant effect of emission mitigation on firm value. Misani et al. (2011) study the relationship between the firm carbon intensity and its market value by analyzing a worldwide sample of 164 firms that have disclosed their greenhouse gas emissions through the CDP in 2006–2008. Similar to Aggarwal and Dow (2011) they also measure firms' organizational responsiveness to climate change by using the qualitative items in the CDP questionnaire. Misani et al. (2011) conclude that carbon intensity is negatively associated with firm market value and that organizational responsiveness to climate change moderates this negative relationship. They suggest that firms that strive to define environmental strategies to reduce their greenhouse gas emissions protect themselves against the negative valuation that investors assign to high polluting companies. These results are in line with Kim and Lyon (2011) who demonstrate that investors view CDP participants as better prepared to cope with future environmental regulation. These studies provide evidence for the resource-based view which states that proactive firms have a competitive advantage over their less green competitors.

Matsumura et al. (2011) investigate the relationship between carbon emissions and firm value for S&P 500 firms disclosing their carbon emissions to the Carbon Disclosure Project over the period 2006–2008. Using a modified variant of the Ohlson valuation model, Matsumura et al. (2011) find a negative association between carbon emission levels and firm value, contingent upon managers' decisions to disclose this non-financial information in the first place. This negative relationship is most pronounced for the high carbon-intensive companies within their sample. Further, Matsumura et al. (2011) indicate that their results involve, on average, a penalty of $202 for every additional metric ton[10] of carbon emissions. They note that this penalty is large considering spot carbon prices below $40 per metric and suggest there are also indirect costs associated with carbon emissions, including potential litigation costs, remediation costs, and loss of reputation, which together add to the total carbon emissions cost. Griffin et al. (2011) also employ a modified Ohlson-type valuation model to assess the effect of greenhouse gas emissions on firm value. They use CDP data on US firms from the S&P 500 over the period 2006–2009 and Canadian firms from the TSE 200 over the period 2005–2009. To control for the self selection bias they estimate carbon emissions for non-disclosures based on the data provided by firms that do disclose. Similar

[10] Greenhouse gas emissions are typically expressed in metric tons, an international unit of measurement equivalent to approximately 2200 pounds (EPA, 2011).

to Matsumura et al. (2011) they conclude that greenhouse gas emissions are negatively associated to firm value and this negative relation is more pronounced for high carbon-intensive firms. This finding is in line with the stakeholder theory as high polluting firms face higher pressure from environmental groups, compared to less polluting firms. Additionally pollution measures capture the exposure of high pollution firms to future environmental liabilities.

Lee et al. (2013) use the event study methodology to investigate market responses to the release of Korean CDP data for 16 firms in 2008 and 50 firms in 2009, all from the KRK 100 index. In addition they examine the moderating effect of frequent carbon communication on the relationship between carbon disclosure and shareholder value by examining a number of articles from 22 Korean newspapers. Their results show that the market responds negatively to firms' carbon emissions disclosure and that a firm can mitigate this negative shock by releasing its carbon news periodically through the media in advance of its carbon disclosure. This last result is not very surprising as the information effect of CDP disclosure is small when firms provide carbon information beforehand.

The CDP data is however not without its limitations. Because the CDP is a voluntary program, firms can respond as they see fit. They can provide all or some of the requested information, or they can decline to participate (Knox-Hayes and Levy, 2011).

In concluding this section, we note that the impact of pollution data disclosed by the Carbon Disclosure Project, on financial performance, does not yield different results compared to pollution data provided within a mandatory context: A firm's polluting level is negatively correlated to its financial performance. There is however some evidence that the negative link between pollution and firm value can be mitigated by voluntarily organizational responsiveness to climate change which is in line with the stakeholders and resource-based view. Through participation in VEPs firms are able to mitigate environmental pressure imposed by various stakeholders.

B. OTHER VOLUNTARY APPROACHES

The US EPA is the largest sponsor of US-based voluntarily environmental programs. By the end of the 1990s, about 13000 firms were participating in EPA-sponsored VEPs (Mazurek, 2002). The 33/50 program was the first voluntary program established by the US EPA. It was established in 1991 with the goal of reducing the aggregate releases of 17 toxic chemicals by 33 percent by 1992 and by 50 percent by 1995, relative to the level in 1988. Firms had the flexibility in the extent of reduction they achieved and in the methods they chose to reduce their releases (EPA, 1999).

Khanna and Damon (1998) use a sample of 123 firms over the period 1991–1993 to examine the impact of the 33/50 program on firms' return on investment (ROI) and expected long run profitability of firms, measured as the excess of market value over the book value of

assets normalized by sales. Controlling for sample selection bias, they conclude that the 33/50 program has a significant, negative effect on the ROI but its impact on the long-term profitability is significantly positive. These results imply that the costs of pollution investments were not offset in the short run by improvement in consumer goodwill and improvements in production efficiency. In the long run however, investors expect that the pollution control efforts will improve the firm's profitability.

Keele and DeHart (2011) use the event study approach to assess how the stocks of publicly traded companies respond to the announcement of their partnership with EPA's Climate Leaders program, a VEP established in 2002. Each US based company that voluntarily joins this program commits to fulfilling a corporate-wide greenhouse gas inventory and to working with EPA to set a corporate emission reduction target. Using a sample of 29 firms, Keele and DeHart (2011) show that the stocks earn an average non-significant positive abnormal return of 0.56% on the day of the announcement, although the cumulative abnormal returns for the stock prices of the firms for two of the three event windows showed statistically significant negative returns. These results suggest that the firms' public announcements of joining the USEPA Climate Leaders did not have a positive impact on stock performance. Fisher-Vanden and Thorburn (2011) also examine the announcement effect of firms joining Climate Leaders. Based on a larger sample of 74 firms their results reveal that companies announcing membership in EPA's Climate Leaders experience significantly negative abnormal returns. Further Fisher-Vanden and Thorburn (2011) also examine 20 announcements of firms joining Coalition for Environmentally Responsible Economies (CERES), a program involving more general environmental commitments, and show that these announcements are associated with insignificant abnormal returns. Finally, given the significantly negative impact on stock price from Climate Leaders membership, they analyse why firms would join this VEP by conducting a probit analysis. The results show that firms with a higher number of shareholder resolutions directed at climate change are more likely to be members of the Climate Leaders program. Controlling for these resolutions, they also find that firms with weak corporate governance structures are more likely Climate Leaders members. Fisher-Vanden and Thorburn (2011) conclude that firms are joining the Climate Leaders program either because they are facing institutional pressures to do so, or because managers face less shareholder oversight, allowing them the possibility to join voluntarily environmental programs. From the agency theory perspective the second explanation is more plausible as it would explain the negative abnormal returns following the announcement to join Climate Leaders by arguing that managers join Climate Leaders to further their own agendas, at the expense of shareholders.

A voluntary environmental program that does not fall under the umbrella of the EPA is the Chicago Climate Exchange (CCX). The CCX was established in 2003 and launched trading operations of the first cap and trade system in North America that made voluntary but legally binding commitments to reduce six different types of greenhouse gas emissions. Market participants included major corporations, utilities and financial institutions with activities in

all 50 United States, 8 Canadian provinces and 16 countries. The total program baseline covered approximately 700 million metric tons of carbon dioxide (CO2), equal to roughly one-third the size of Europe's cap and trade program (Chicago Climate Exchange, 2011). Gans and Hintermann (2013) analyse the stock return behaviour of member firms of the Chicago Climate Exchange (CCX) on a monthly basis. They base their study on a difference-in-difference framework. To control for self-selection bias into the voluntary program, they construct control groups of non-member firms based on propensity score matching. The authors find positive and statistically significant excess returns after firms announce their decision to join CCX. These findings are not in line with the results of Fisher-Vanden and Thorburn (2011) who find that firms announcing membership in the VEP Climate Leaders experience significantly negative abnormal returns. Gans and Hintermann (2013) additionally investigate the financial impact of the Waxman-Markey Bill on members of the CCX. The Waxman-Markey Bill proposed, among other things, the introduction of a cap-and-trade system which would regulate the emission of greenhouse gases in the United States. Under the bill, over the next 40 years carbon emissions would be increasingly decreased to 83% compared 2005 levels. Gans and Hintermann (2013) argue that this bill raised the likelihood of a mandatory cap-and-trade system being instituted in the medium term. They find that the passing of the Waxman-Markey climate bill leads to positive and statistically significant excess returns for CCX member firms relative to non-member firms, implying that firms who had gained experience in the voluntary market are rewarded for being prepared for future regulation. This finding is similar to the study of Kim and Lyon (2011) who find that companies' CDP participation increased shareholder value following Russia's ratification of the Kyoto Protocol.

Overall, studies that examine the relationship between market value and partnerships with VEP's find mixed results. These inconclusive results could be attributable to different beliefs of investors about the benefit of membership or different goals altogether. If joining a VEP introduces an agency problem between managers and shareholders, announcing participation should be associated with negative abnormal returns. On the other hand, following the stakeholders perspective, joining VEPs should be rewarded if VEPs are perceived as helpful tools to satisfy stakeholders' environmental demands. Firms participating in voluntary initiatives should therefore communicate the benefits of VEPs clearly with their investors to avoid any agency problems.

C. DISCLOSURE OF GREEN PERFORMANCE

The link between green performance and firm performance, within the framework of mandatory or voluntary environmental programs, has been subject to a large amount of research, as reported in previous sections. Another set of studies analyses the impact of corporate green performance disclosures, published by the company itself or by non-profit organizations, on market value.

As stakeholder theory suggests that firms must satisfy several groups (e.g., the government, non-governmental organizations, employees) that have some interest or in a firm, It can be worthwhile for firms to report on environmental efforts because otherwise these stakeholders could withdraw their support (Ziegler et al., 2011). Furthermore, from the resource based view perspective, a good reputation due to the disclosure of corporate environmental measures is an example of an intangible resource. Overall, the stakeholder and resource based view theory suggest that the link between disclosed corporate environmental efforts should be positive.

However, it can also be argued that the disclosure of corporate environmental activities is not a result of proactivity. In contrast, the disclosure of corporate environmental efforts could be an answer to institutional pressures due to the increasing discussion about climate change (Ziegler et al., 2011). In this case, pollution abatement activities could be a reactive strategy so that the demanded investments lead to unexpected costs (e.g., King and Lenox, 2001). Following the agency cost theory this cost argument is considered the standard argument for a negative relationship between corporate environmental performance and financial performance. As a consequence the disclosure of corporate environmental efforts should have a negative impact on a firm's financial performance.

An example of a non-profit organization disclosing corporate green performance information is Climate Counts. Its goal is putting pressure on corporations into reducing contributions to climate change. Climate Counts scores the world's largest companies on their climate impact to spur corporate climate responsibility and conscious consumption (Climate Counts, 2013). Beatty and Shimshack (2010) explore the capital market impact of the disclosure of Climate Counts' scores in June 2007. They find, by conducting an event study with 47 observations, that the release of climate ratings had a significant impact on stock prices. This result is primarily driven by penalties to firms receiving poor climate performance ratings while this study does not provide significant evidence that good ratings are associated with positive abnormal returns.

Griffin and Sun (2013) analyse the announcement effect of firms' voluntary disclosures about greenhouse gas emissions made through the Corporate Social Responsibility newswire service (CSRwire), a digital media platform that claims to be the global leader in climate change disclosure. For a sample of 172 disclosures by 84 US listed companies over 2000–2010, they document that the voluntary green disclosure provides shareholders with positive returns. Ziegler et al. (2011) analyse the relationship between disclosed corporate responses to climate change and stock performance. They use a sample of European and US firms across the time period of 2001 to 2006 and argue that the awareness of climate change and the stringency of climate policy were generally higher in Europe compared to the USA. In contrast to studies using long-term firm performance indicators or short-run event studies, Ziegler et al. (2011) examine the average stock performance of portfolios that differ in their disclosure practices. In order to estimate the corresponding risk-adjusted returns, they apply the four-factor model according to Carhart (1997) in addition to the one-factor model based on the CAPM. The

results document that a trading strategy which consists of buying stocks of corporations disclosing responses to climate change and selling stock of corporations with no disclosures has become more worthwhile over time in Europe and has particularly been rewarded in the period of 2004 to 2006. The authors conclude that the stock performance of firms with a higher level of disclosed responses to climate change is slightly more positive in regions and periods with a higher institutional pressure with respect to global warming and thus a more stringent general or sectoral climate policy regime than in regions and periods with weaker climate policy. Another finding is that the relationship between disclosed corporate responses to climate change and stock performance is positive for energy firms. Ziegler et al. (2011) argue that their findings support the stakeholder theory as stakeholders have a bigger appetite for environmental disclosures in regions with higher climate change awareness. In addition energy firms are more severely observed, for example, by non-governmental organizations. As a consequence, good relationships with stakeholders are more important for this group of firms (e.g., Sprengel and Busch, 2011). Altogether, previous discussed studies document that the disclosure of voluntary environmental efforts are positively valued by investors as they address stakeholders' environmental demands and obtain goodwill.

V. CONCLUSION, IMPLICATIONS AND FUTURE RESEARCH

The aim of this article was to review and discuss the academic literature regarding the link between pollution performance and financial performance. We address the question whether the link between environmental and financial performance is driven by environmental regulation or by disclosure efforts through voluntary environmental initiatives.

According to early agency arguments firms that invest in pollution control will incur costs that outweigh their financial benefits. As a consequence, corporate environmental investments can lead to reduced profits, decreased firm values, or competitive disadvantage and therefore result in lower profit expectations of investors (Aupperle et al., 1985). Subsequently, researches have challenged this view indicating that "it pays to be green". Argued arguments from a stakeholders and resource-based view perspective include obtaining goodwill and standing with critical stakeholders, cost efficiencies associated with innovation, gaining a competitive advantage and reducing risks of future mandatory regulation. Following the stakeholder perspective, environmental regulation seems less relevant as stakeholder will enforce environmental responsibility upon the company without legislation imposed.

Overall, this review supports the positive link between corporate environmental performance and financial performance. The literature provides quite consistent evidence of a negative relationship between firms' emissions, both within a regulated and voluntary framework, and financial performance. Additionally, several studies found a positive association between pollution reduction and firm value.

Studies that evaluate the link between market value and partnerships with voluntary environmental programs are inconclusive. These inconclusive results could be attributable to different beliefs of investors about the benefit of participation in VEP's. Several studies find however that voluntary environmental initiatives mitigate the negative effects of pollution on firm value in the light of stricter environmental regulations.

The positive link between, pollution reduction and firm value and the mitigating effect of voluntary environmental initiatives and disclosures provide evidence for the stakeholder theory and resourced-based-view theory suggesting that firms can improve their financial performance by satisfying stakeholders' demands and implementing a proactive strategy. It is the stakeholders therefore who, enforced by enhancing environmental corporate disclosure, truly force firms to implement environmental practices, while governmental pressures are less relevant.

This review implies that managers do not face a trade off between environmental and financial performance. Thus, the reduction of polluting emissions as well as other environmental performance is an appropriate business strategy that does not conflict with firms' economic incentives. Therefore, managers are recommended to reduce polluting emissions to enhance their economic performance and disclose these initiatives effectively to their stakeholders. Further, this review shows that a proactive approach towards environmental requirements instead of meeting compliance at minimum costs could create a sustainable advantage. Firms engaging in voluntary initiatives should however communicate the benefits of participation clearly with their investors to avoid any agency problems. In fact firms should include environmental performance as an integral part of corporate strategy, allowing managers the time and resources they need to manage the environmental challenges. A clear proactive environmental strategy should not only guide the development of competencies but also shape the firm's relationship with employees, suppliers, customers, policy makers, and all other stakeholders (Hart, 1997). However, according to Clarkson et al. (2011c) only firms with sufficient financial resources and management capabilities can pursue a proactive environmental strategy. Given these resource constraints, policy makers should provide firms with incentives to improve their environmental performance. For instance, higher tax benefits associated with green investments or market-based mechanism such as emission trading schemes. Such schemes also provide the public with benchmarks against which good and bad environmental performance can be defined. In addition, environmental regulation needs to encourage the participation in VEPs rather than to penalize polluters. Technical assistance provided by such programs can assist firms in understanding and identifying technical solutions that are needed to address their environmental challenges. Furthermore policies should provide channels through which superior environmental performance can be disclosed. Public recognition of superior environmental performance can be a substantial incentive as such recognition could lead to economic gains in the form of stakeholders' goodwill.

Further, these findings do not necessarily imply that investors can use information about a firm's environmental strategy to earn abnormal returns, they do suggest that investors and analysts should consider a firm's environmental performance when forming investment strategies. This review shows that firms following a more proactive environmental strategy, satisfying the demands of various stakeholders, are more likely to attract investors to buy their stocks and enhance their share returns ultimately.

In terms of future research, it would be interesting to investigate, following Ziegler et al. (2011), how institutional pressure with respect to global warming contributes to the relationship between environmental and financial performance. Further, the research on the relationship between carbon performance and firm performance, within the framework of cap-and-trade programs, such as the EU ETS or the Australian ETS, is rather scarce. More research is needed to understand the impact of emission trading on firm performance as it is not clear whether the impact is driven by a reputation or a value effect, or a combination of the two.

Research on the link between pollution and firm performance has been based on pollution data such as that provided by the TRI database, only representing certain aspects of environmental performance. In this sense the relationship between environmental and firm performance has only partially been examined. In terms of future research, it would be interesting to discover the full picture of the environmental performance -firm performance link. However, this task is far from easy due to lack of complete data on firms' environmental performance. Next, the majority of studies summarized in this review examine the relationship between environmental and financial performance from the perspective of market-based measures of firm performance. There is considerably less research, especially in a non-US context that has focused on the link between a firm's environmental performance and its cost of capital. Finally, further research is recommended to examine how firm characteristics affect the link between environmental performance and firm performance. The relationship between firm characteristics and environmental performance is likely to be complex.

REFERENCES

Aggarwal, R. and Dow, S. (2011), *Greenhouse gas emissions mitigation and firm value: a study of large North-American and European firms*. working paper.

Alberini, A. and Segerson, K. (2002), "Assessing voluntary programs to improve environmental quality", *Environmental Resources Economics*, 22: 157–184.

Anger, N. and Oberndorfer, U. (2008), "Firm performance and employment in the EU emissions trading scheme: An empirical assessment for Germany", *Energy Policy*, 36: 12–22.

Annandale, D., Morrison-Saunders, A. and Bouma, G. (2004.), "The impact of voluntary environmental protection instruments on company environmental performance", *Business Strategy and the Environment*, 13(1): 1–12.

Ataulah, A., Higson, A. and Tippett, M. (2006), "Real (adaptation) options and the valuation of equity: Some empirical evidence", *Abacus*, 42(2).

Aupperle, K., Carrol, A. and Hatfield, J. (1985), "An empirical examination of the relationship between corporate social responsibility and profitability", *Academy of Management Journal*, 28: 446–463.

Bao, C., Lu, Y. and Shang, J. (2004), "Framework and operational procedure for implementing strategic environmental assessment in China", *Environmental Impact Assessment Review*, 24(1): 27–46.

Barney, J. B. (2001), "Is the resource-based "view" a useful perspective for strategic management research? Yes", *The Academy of Management Review*, 26(1): 41–56.

Bates, G. M. (2010) *Environmental Law in Australia*. Sydney: Butterworths.

Beatty, T. K. M. and Shimshack, J. P. (2010), "The impact of climate change information: new evidence from the stock market", *The Berkely Electronic Journals of Economic Analysis and Policy*, 10(1): 1635–1682.

Blacconiere, W. G. and Patten, D. M. (1994), "Environmental disclosures, regulatory costs, and changes in firm value", *Journal of Accounting and Economics*, 18: 357–377.

Blacconiere, W. G. and Northcut, W. D. (1997), "Environmental information and market reaction to environmental legislation", *Journal of Accounting, Auditing and Finance*, 149–178.

Burtraw, D., Evans, D. A., Krupnick, A., Palmer, K. and Toth, R. (2005), "Economics of pollution trading for SO2 and NOx", *Annual Review of Environmental Resources*, 30: 253–289.

Calvin, K., Clarke, L., Krey, V., Blanford, G., Jiang, K., Kainuma, M., Kriegler, E., Luderer, G. and Shukla, P. R. (2012), "The role of Asia in mitigating climate change: Results from the Asia modeling exercise", *Energy Economics,* 34: 251–260.

Carbon Disclosure Project. (2013), *The Carbon Disclosure Project Homepage*. Retrieved October 13, 2013, from https://www.cdproject.net/en-US/Pages/HomePage.aspx.

Carhart, M. (1997), "On persistence in Mutual Fund Performance", *Journal of Finance*, 52: 57–82.

Carriker, R.R. (1996), "Federal Environmental Policy: A Summary Overview", *Journal of Agricultural and Applied Economics*, 28: 99–107.

Caves, R. E. and Porter, M. E. (1977), "From entry barriers to mobility barriers: Conjectural decisions and contrived deterrence to new competition", *Quarterly Journal of Economics*, 91: 241–262.

Chapple, L., Clarkson, P. M. and Gold, D. L. (2013), "The cost of carbon: Capital market effects of the proposed emission trading scheme (ETS)", *ABACUS*, 49(1).

Chicago Climate Exchange (2011), CCX Fact sheet, Chicago.

Clarkson, M. (1995), "A stakeholder framework for analyzing and evaluating corporate social performance", *Academy of Management Review*, 20(1): 92–117.

Clarkson, P. M. and Li, Y. (2004) "The market valuation of environmental capital expenditures by pulp and paper companies", *The Accounting Review*, 79(2): 329–353.

Clarkson, P. M., Fang, X. and Li, Y. (2011a), *The relevance of environmental disclosures for investors and other stakeholder groups: are such disclosures incrementally informative*, working paper.

Clarkson, P. M., Li, Y., Richardson, G. and Vasvari, F. P. (2011b), "Does it really pay to be green? Determinants and consequences of proactive strategies", *Journal of Accounting Public Policy*, 30: 122–144.

Clarkson, P. M., Overell, M. B. and Chapple, L. (2011c), "Environmental reporting and its relation to corporate environmental performance", *Abacus*, 47(1): 27–60.

Clarkson, P. M., Li, Y., Pinnuck, M. and Richardson, G. (2014), "The value relevance of greenhouse gas emissions under the European Union Carbon Emissions Trading Scheme", *European Accounting review*, forthcoming.

Climate Counts (2013) *Climate Counts, about us*. Retrieved October 25, 2013, from, www.climatecounts.org/about.php

Cohen, M. A., Fenn, S. A. and Konar, S. (1997), *Environmental and financial performance: are they related?*, working paper.

Connors, E. and Silva-Gao, L. (2008), *The impact of environmental risk on the cost of equity capital: evidence from the toxic release inventory*, working paper.

Cordeiro, J. J. and Sarkis, J. (1997), "Environmental Proactivism and firm performance: Evidence from security analyst earnings forecasts", *Business Strategy and the Environment*, 6: 104–114.

Cormier, D., Magnan, M. and Morard, B. (1993), "The impact of corporate pollution on market valuation: some empirical evidence", *Ecological Economics*, 8: 135–155.

Darnall, N., and Carmin, J. (2005), "Greener and cleaner? The signalling accuracy of US voluntary environmental programs", *Policy Sciences*, 38: 71–90.

Derwall, J., Guenster, N., Bauer, R. and Koedijk, K. (2005), "The Eco-Efficiency Premium Puzzle", *Financial Analysts Journal*, 61 (2): 51–63.

European Commission (2013), The EU Emission Trading System (EU ETS), Brussels.

Environmental Protection Agency (1999), 33/50 Program. The Final Record, Washington.

Environmental Protection Agency (2011), Greenhouse gas emissions from a passenger vehicle, Washington.

Fisher-Vanden, K. and Thorburn, K. S. (2011), "Voluntary corporate environmental initiatives and shareholder wealth", *Journal of Environmental Economics and Management*, 62: 430–445.

Fornaro, J. M., Winkelman, K. A. and Glodstein, D. (2009), "Accounting for emissions", Journal of Accountancy, July: 40–45.

Freeman, E. (1984), *Strategic Management: a stakeholder approach*. New Jersey.

Freedman, M. and Jaggi, B. (2009), "Global warming and corporate disclosures: A comparative analysis of companies from the European Union, Japan, and Canada", *Advances in environmental accounting and management*, 4: 129–160.

Friedman, M. (1970), The social responsibility of business is to increase its profits, *New York Times*. Retrieved www.colorado.edu/studentgroups/libertarians/issues/friedman-soc-resp-business.html.

Gans, W. and Hintermann, B. (2013), "Market effects of voluntary climate action by firms: Evidence from the Chicago Climate Exchange", *Environmental and Resource Economics*, 55(2): 291–308.

González-Benito, J. and González-Benito, O. (2006), "The role of stakeholder pressure and managerial values in the implementation of environmental logistics practices", *International Journal of Production Research*, 44(7): 1353–1373.

Gray, W.B. and Shadbegain, R.J. (1995), *Pollution Abatement Costs, Regulation, and Plant-Level Productivity*, working paper.

Griffin, P., Lont, D. and Sun, Y. (2011), *The relevance of investors of greenhouse gas emission disclosures. working paper*, working paper.

Griffin, P. and Sun, Y. (2013), "Going green: market reactions to CSR newswire releases". *Journal of Accounting and Public Policy*, 32(2): 93–113.

Guenster, N., Bauer, R., Derwall, J. and Koedijk, K. (2011), "The economic value of corporate eco-efficiency", *European Financial Management*, 17(4): 679–704.

Hamilton, J. T. (1995), "Pollution as news: Media and stock market reactions to the toxics release inventory data", *Journal of Environmental Economics and Management*, 28(1): 98–113.

Hart, S. L. (1997), "Beyond greening. Strategies for a sustainable world", *Harvard Business Review*, 24(1): 66–76.

Hart, S. L. and Ahuja, G. (1996), "Does it pay to be green? An empirical examination of the relationship between emission reduction and firm performance", *Business Strategy and the Environment*, 5: 30–37.

Heckman, J. J. (1979), "Sample selection bias as a specification error", *Econometrica*, 47(1): 53–161.

Hrasky, S. (2011), "Carbon footprints and legitimation strategies: symbolims or action?", *Accounting, Auditing & Accountability*, 25(1): 174–198.

Hughes, K. E. (2000), "The value relevance of nonfinancial measures of air pollution in the electric utility industry", *The Accounting Review* 75: 1199–1214.

Ilinitch, A., Soderstrom, N. and Thomas, T. (1998), "Measuring corporate environmental performance", *Journal of Accounting and Public Policy*, 17(383–408).

Jaggi, B. and Freedman, M. (1992), "An examination of the impact of pollution performance on economic and market performance: pulp and paper firms", *Journal of Business Finance and Accounting*, 19(5).

Johnston, D., Sefcik, S. E. and Soderstrom, N. S. (2008), "The value relevance of greenhouse gas emissions allowances: an exploratory study in the related United States SO_2 market", *European Accounting Review*, 17(4): 747–764.

Keele, D. M. and DeHart, S. (2011), "Partners of USEPA Climate Leaders: an event study on stock performance", *Business Strategy and the Environment*, 20: 485–497.

Khanna, M. and Damon, L. A. (1998), "EPA's voluntary 33/50 program: Impact on toxic releases and economic performance of firms", *Journal of Environmental Economics and Management*, 37: 1–25.

Khanna, M., Quimio, W. and Bojilova, D. (1998), "Toxics release information: A policy tool for environmental protection", *Journal of Environmental Economics and Management*, 36(243–266).

Kim, E. and Lyon, T. P. (2011), "When does institutional investor activism pay? The Carbon Disclosure Project", *The B.E. Journal of Economic Analysis & Policy*, 11(1).

King, A. and Lenox, M. (2001), "Does it really pay to be green", *The Journal of Industrial Ecology*, 5(1): 105–116.

Knox-Hayes, J. and Levy, D. (2011), "The politics of carbon disclosure as climate governance", *Strategic Organization*, 9(1).

Konar, S. and Cohen, M. A. (2001), "Does the market value environmental performance", *The Review of Economics and Statistics*, 83(2): 281–289.

Lee, S., Park, Y. and Klassen, R. D. (2013), "Market responses to firms' voluntary climate change disclosure and carbon communication", *Corporate Social Responsibility and Environmental Management*.

Matsumura, E. M., Prakash, R. and Vera-Munoz, S. C. (2011), *Carbon emissions and firm value*, working paper.

Mazurek, J. (2002) *Government Sponsored Voluntary Programs: An initial Survey. In T. D. a. P. C. Stern (Ed.), New Tools for Environmental Protection: Education, Information and Voluntary Measures.* Washington DC: National Academy Press.

McWilliams, A., Siegel, D. and Wright, P. (2001), "Corporate social responsibility: stratgeic implications", *Journal of Management Studies*, 43: 1–18.

Miclaus, P. G., Lupu, R., Dumitrescu, S. A. and Bobirca, A. (2008), "Futures Market using Event-Study Methodology", *International Journal of Energy and Environment*, 2(2).

Miles, M. P. and Covin, G. (2000), "Environmental marketing: A source of reputational, competitive and financial advantage", *Journal of Business Ethics*, 23: 299–311.

Misani, N., Pogutz, S. and Russo, A. (2011), *Investigating the relationship between firm carbon intensity and financial performance: the role of organizational responsiveness*, working paper.

Mnif, D. and Davison, M. (2012), *EU ETS futures spread analysis and pricing contingent claims*, working paper.

Nishitani, K., Kaneko, S., Komatsu, S. and Hidemichi, F. (2011), *Firms reduction of greenhouse gas emissions and economic performance: analyzing effects through demand and productivity*, working paper.

Nishitani, K. and Kokubu, K. (2011), "Why does the reduction of greenhouse gas emissions enchance firm value? The case of manufacturing firm", *Business Strategy and the Environment*.

Ohlson, J. (1995), "Book value, and dividens in security valuation", *Contemporary Accounting Research*: 661–687.

Palmer, K., Oates, W. E. and Portney, P. R. (1995), "Tightening environmental standards: The benefit-cost or the no-cost paradigm?", *Journal of Economic Perspectives*, 9(4): 119–132.

Patten, D. (2002), "The relation between environmental performance and environmental disclosure: a research note", *Accounting, Organizations and Society*, 27(8): 763–773.

Porter, M. E. and Van der Linde, C. (1995), "Green and competitive: ending the stalemate", *Harvard Business Review*, 73(5): 120–134.

Prakash, A. and Potoski, M. (2012), "Voluntary environmental programs: A comparative perspective, *Journal of Policy Analysis and Management,* 31(1): 123–138.

Ramiah, V., Martin, B. and Moosa, I. (2013), "How does the stock market react to the announcement of green policies", *Journal of Banking and Finance*, 37: 1747–1758.

Rassier, D. G. and Earnhart, D. (2010), "Does the Porter hypothesis explain expected future financial performance? The effect of clean water regulation on chemical manufacturing firms", *Environmental and Resource Economics*, 45(3): 353–377.

Revesz, R. L. and Stavins, R. N. (2003) *Environmental Law and Policy, The Handbook of Law and Economics*. Amsterdam: North -Holland/Elsevier Science.

Rivera-Camino, J. (2007), "Re-evaluating green marketing strategy: a stakeholder perspective", *European Journal of Marketing*, 41(11/12): 1328–1358.

Ruf, B. M., Muralidhar, K., Brown, R. M., Janny, J. J. and Paul, K. (2001), "An empirical investigation of the relationship between change in corporate social performance and financial performance: a stakeholder theory perspective", *Journal of Business Ethics*, 32(2): 143–156.

Russo, M. and Fouts, P. (1997), "A resource-based perspective on corporate environmental performance and profitabiliy", *Academy of Management Journal*, 40: 534–559.

Saha, S. and Mohr, R. D. (2013), "Media attention and the Toxics Release Inventory", *Ecological Economics*, 93.

Sarkis, J. and Cordeiro, J. J. (2001), "An empirical evaluation of environmental efficiencies and firm performance: pollution prevention versus end-of-pipe practice", *European Journal of Operational Research*, 135: 102–113.

Schmidt, P. and Werner, T. (2012), *Verified emissions and stock prices: Is there a link? An empirical analysis of the European Emission Trading Scheme*, working paper.

Shane, P. B. and Spicer, B. H. (1983), "Market response to environmental information produced outside the firm", *The Accounting Review*, 58(3).

Sprengel, D. C. and Busch, T. (2011), "Stakeholder engagement and environmental strategy, the case of climate change mitigation policy", *Urban Affairs Review*, 47(3): 433–457.

Stanny, E. and Ely, K. (2008), "Corporate environmental disclosures about the effects of climate change", *Corporate Social Responsibility and Environmental Management*, 15(6): 338–348.

Stavins, R. (2007), *A U.S. Cap-and-Trade System to Address Global Climate Change*, working paper.

Stewart, R. (1993), "Environmental regulation and international competitiveness", *Yale Law Journal*, 102: 2039–2106.

Tietenberg, T. H. (1990), "Economic instruments for environmental regulation", *Oxford Review of Economic Policy*, 6(1): 17–33.

Toffel, M. W. and Marshall, J. D. (2004), "Improving environmental performance assessment: A comparative analysis of weighting methods used to evaluate chemical release inventories", *Journal of Industrial Ecology*, 8(1–2): 143–172.

Tsang, S. and Kolk, A. (2010), "The evolution of Chinese policies and governance structures on environment, energy and climate", *Environmental Policy and Governance*, 20(3): 180–196.

Videras, J. and Alberini, A. (2000), "The appeal of voluntary environmental programs: which firms participate and why?", *Contemporary Economic Policy*, 18(4): 449–461.

Walley, N. and Whitehead, B. (1994), "It's not easy being green", *Harvard Business Review*, 72: 46–52.

Wayne, B. G. and Shadbegian, R. J. (1995), *Pollution abatement costs, regulation and plant-level productivity*, working paper.

Wernerfelt, B. (1984), "A resource based view of the firm", *Strategic Management Journal*, 5: 171–180.

Wood, D. and Ross, D. G. (2006), "Environmental social controls and capital investments:Australian evidence", *Accounting and Finance*, 46: 677–695.

Ye, D., Liu, S. and Kong, D. (2013), "Do efforts on energy saving enhance firm values? Evidence from China's stock market", *Energy economics*, 40: 360–369.

Yongqing, L., Eddie, I. and Jingui, L. (2013), "The impact of carbon emissions on asset values and operating cash flows: Evidence from Australian listed firms ", *Journal of modern Accounting and Auditing*, 9(1): 94–111.

Ziegler, A., Bush, T. and Hoffman, V. H. (2011), "Disclosed corporate responses to climate change and stock performance: An international empirical analysis", *Energy Economics*, 33: 1283–1294.

Ziegler, A., Schröder, M. and Rennings, K. (2007),"The effect of environmental and social performance on the stock performance of European corporations", *European Association of Environmental and Resource Economists*, 37(4): 661–680.

DISCLOSURE OF NON-FINANCIAL INFORMATION: RELEVANT TO FINANCIAL ANALYSTS?

Raf Orens* and Nadine Lybaert**

Abstract

The decline in the relevance of financial statement information to value firms leads to calls from organizational stakeholders to convey non-financial information in order to be able to judge firms' financial performance and value. This literature review aims to report extant literature findings on the use of corporate non-financial information by sell-side financial analysts, the information intermediaries between corporate management and investors. Prior studies highlight that financial analysts rely upon corporate non-financial information in their firm assessments. When firms to a larger extent disseminate non-financial information, financial analysts are able to submit more accurate earnings forecasts and the consensus among financial analysts' earnings estimates is larger. However, the literature review also illustrates that financial analysts ascribe more weight to particular types of non-financial information. For instance, they consider forward-looking information or strategy and product-related information more relevant in firm valuation compared to intellectual capital information or social and environmental information.

Keywords: financial analysts; non-financial information; voluntary disclosure

JEL codes: M41, M48

* (corresponding author) KU Leuven, Faculty of Business and Economics @ Thomas More Antwerpen, Korte Nieuwstraat 33, 2000 Antwerpen, Belgium, E-mail: raf.orens@kuleuven.be.

** Hasselt University, KIZOK, Agoralaan Gebouw D, 3590 Diepenbeek, Belgium, E-mail: nadine.lybaert@uhasselt.be.

I. INTRODUCTION

This article aims to review prior literature on the relevance of non-financial information in the decision-making process of financial analysts. Information conveyed by firms is relevant when financial analysts rely on it in their equity valuation or in their forecasting work (Cormier and Magnan, 2013). The globalization, technological evolutions, and the transition towards a knowledge economy increase the usefulness of non-financial information in judging firm value in addition to financial information (Arvidsson, 2012; IFAC, 2013). The decline in relevance of financial information in explaining a firm's value (Lev and Zarowin, 1999; Francis and Schipper, 1999; Dontoh et al., 2004; Liang and Yao, 2005; Hail, 2013), leads to the recognition that financial statement information is insufficient to satisfy the information needs of stakeholders to assess firms' performance. Stakeholders put pressure on firms to report non-financial information about their strategy, their investments in research and development or their customer satisfaction levels, in order to judge firm performance and to predict future earnings. Since regulatory requirements related to non-financial information are almost absent (IFAC, 2013), firms have to report such information voluntarily.

The current literature review focuses on studies investigating the use of non-financial information by sell-side financial analysts[1], as these stakeholders are important users of corporate information (Luo et al., 2014). Sell-side financial analysts are employed by brokering firms, investment banks or research firms to assess the performance of listed firms and to disseminate corporate information, earnings estimates and stock recommendations to their clients, such as retail and institutional investors (Beyer et al., 2010; Pinho et al., 2013). These investors rely upon the financial analysts' reports and recommendations to make investment decisions (Fogarty and Rogers, 2005; Johansson, 2007; Groysberg et al., 2008). Sell-side analysts cover a limited number of listed firms and are often industry specialists.

Financial analysts perform two different and important roles in capital markets. First, financial analysts are intermediaries collecting information from firm management and relaying this information to investors (Lang and Lundholm, 1996; Barker, 1998; Ivković and Jegadeesh, 2004; Cheng et al., 2006, Chen et al., 2010). Analysts add value to investors by transforming public and private information into earnings forecasts and stock recommendations which are used by investors to make investment decisions (Hong et al., 2000; Elgers et al., 2001; Chen et al., 2010). Especially with regard to non-financial information, financial analysts add value to investors by translating the bulk of non-financial information disclosed by firms into comprehensive information to investors. Although investors' needs of non-financial information increased over time, investors have difficulties to interpret the value and the earnings effects of non-financial disclosures (Maines and McDaniel, 2000;

[1] Buy-side financial analysts are hardly examined in prior literature since they rely upon information from sell-side financial analysts to make decisions. These analysts typically do not provide detailed assessments and forecasts since they cover a much larger number of firms in comparison to sell-side financial analysts.

Maines et al., 2002; Hoff and Wood, 2008). In addition, no uniform reporting format exists for non-financial disclosures creating differences in the presentation of this information across firms (Simpson, 2010; Eccles et al., 2011), which raises the complexity in analysing non-financial information. The lack of investors' knowledge concerning the valuation impact of non-financial information increases the incentives for financial analysts to clarify how this information impacts firm performance and firm value.

Firm monitoring is a second important function of financial analysts. By assessing firms, analysts are able to attenuate equity agency conflicts between investors or shareholders on the one hand, and corporate management on the other hand (Chung and Jo, 1996; Doukas et al., 2000). The decline in the usefulness of financial statement information to value firms even increases the importance of this monitoring role (Chung et al., 2005; Jiraporn and Gleason, 2007; LaFond and Watts, 2008).

Prior literature uses two approaches to understand the financial analysts' behaviour regarding non-financial information. Some studies indirectly investigate analysts' reliance on corporate non-financial information by relating the extent and quality of non-financial disclosures to properties of analysts' earnings forecasts. Other studies directly address financial analysts' use of non-financial information through questionnaires, interviews, or content-analysing the reports issued by financial analysts.

Based on the literature review, we conclude that financial analysts employ non-financial information in estimating future firm performance and firm value. Firms releasing a larger amount of non-financial information allow financial analysts to report more accurate earnings estimates and to provide less dispersed earnings forecasts. Survey-based evidence and the content analysis of analyst reports also demonstrate the increasing use of non-financial information by financial analysts over time. However, the research findings also document some variation in the types of non-financial information used. In general, financial analysts tend to rely more on forward-looking information, strategy-related information and product-related information. Financial analysts hardly use intellectual capital information or corporate social responsibility information. Potentially, the lower reliability of the latter information elements in predicting firms' cash flows restricts their use.

Presenting extant literature results related to the relevance of non-financial information for financial markets participants is useful to firm managers in order to make decisions regarding their disclosure policy. Since non-financial information leads to proprietary costs, it is important for firms to gauge the benefits and costs of disclosure. The findings are also useful to regulators to understand the non-financial information needs of capital market participants, and hence to potentially consider public reporting for those non-financial information elements which are useful to financial analysts to judge firm performance.

The remainder of the literature review is organized as follows. Section 2 discusses prior literature on the disclosure of non-financial information by firms. Section 3 highlights literature findings on the financial analysts' use of non-financial information. Section 4 provides some concluding remarks and suggestions for further research.

II. THE DISCLOSURE OF NON-FINANCIAL INFORMATION

Prior literature emphasized the increasing importance of non-financial information in judging firms' value over time (Zéghal and Maaloul, 2011; IFAC, 2013). Although financial information remains important in firm valuation and in the decision-making process by investors and other external stakeholders (Cole et al., 2012), this information is of declined value to these stakeholders (Dontoh et al., 2004; Liang and Lao, 2005). Financial statements are inadequate to reflect intangible and other non-financial value drivers, such as customer satisfaction or employee experience (Holder-Webb, 2009; Simnett et al., 2009). The relevance of financial information in firm valuation decreased due to the outdated nature of this information, the discretion employed by managers to estimate financial information and the changing environment in which firms operate. With the latter, the globalization and the growing influence of multinationals, the transition to a knowledge economy, the introduction of new technologies, the financial crisis, the growth in ethical/socially responsible investments[2] or the climate change are considered (Francis et al., 2002, Dhaliwal et al., 2011; Hail, 2013). Financial information hence explains a diminishing part of firms' value, leading to the call from organizational stakeholders for the disclosure of non-financial information. Extant literature already demonstrated that stock prices are affected by the publication of non-financial information, indicating that this type of information is relevant for firm valuations (Berthelot et al., 2003; Fernandez et al., 2011; Matsumoto et al., 2011; Ellis et al., 2012;). This result is not surprising since non-financial information often deals with intangible assets such as employees' knowledge, customer satisfaction or distribution channels which are considered as main value creators for firms (Zéghal and Maaloul, 2011). These assets account for well over half of the market capitalisation of public firms (Lev, 2004).

Despite the general awareness of the increasing importance of non-financial information to judge firm performance, extant literature struggles to define non-financial information consistently. One explanation for this inconsistency is that non-financial information represents or acts on totally different matters, like information with respect to firms' strategy, social responsibility, corporate governance, internal control or risk management (Said et al., 2003; Juntilla et al., 2005). A clear definition of non-financial information is hence lacking. To illustrate, non-financial information is defined as non-accounting information (Amir and Lev, 1996), as non-financial disclosures and metrics including index scores, ratios, counts and other information not presented in the financial statements (Upton, 2001), as information which cannot be directly derived from the financial statements of the company (Cohen et al., 2008), or as all quantitative and qualitative information on the strategy, management and its outcomes in terms of performance or effects, without a direct link with a financial registration system (NIVRA, 2010).

[2] Anecdotal evidence indicates that a firm's reputation and long-term sales can suffer because of poor corporate social responsibility performance. For example, Nike struggled for years and invested a great amount of financial resources and effort to regain its reputation after the 1997 child labor scandal (Dhaliwal et al., 2011).

Although these definitions provide some guidance, they do not lead to an unambiguous assignment. For instance, an earnings forecast issued by firm management – being a metric published outside financial statements – is considered non-financial information according to the Upton's (2001) definition, but following Amir en Lev (1996), this is considered as financial information as an earnings forecast is drawn from financial statements. Some studies equal non-financial information as corporate social responsibility (CSR) information (e.g., Dhaliwal et al., 2011), but this is incorrect since CSR combines financial as well as non-financial information. Other studies equal non-financial information with qualitative information (e.g., IFAC, 2013), but non-financial information could also include quantitative information, and hence this definition does not cover the full package.

Since a clear definition is non-existent, empirical studies focusing on non-financial information do not prefer to start from a definition of non-financial information, but describe this concept using a list of non-financial information elements. The selection of this list of items (or a disclosure index) is based on recommendations issued by accounting standard setters (e.g. FASB, 2001) or federations of accounting professionals (e.g. AICPA, 1994). Some studies develop a self-constructed disclosure index, selecting and grouping non-financial information elements, like for instance in Said et al. (2003), Juntilla et al. (2005), Hoff & Wood (2008), Coram et al. (2011), Eccles et al. (2011).

Despite the potential confusion about the definition of non-financial information, firms respond to the stakeholders' call for the disclosure of non-financial information (Adams et al., 2011). Since ample regulatory requirements exist, firms report non-financial information on a voluntary basis. Unsurprisingly, the growing needs of non-financial information by organizational stakeholders over time lead to an increase in firms' voluntary reporting about non-financial information, both in volume and complexity (Orens and Lybaert, 2007; Cohen et al., 2011). Given that the concept of non-financial information is so broad, and covers a wide range of topics, such as corporate governance disclosures, environmental information, intellectual capital disclosure or strategy and management information, the amount and detail of non-financial information disclosed by firms strongly varies across firms. In addition, firms have the possibility to use various communication platforms to distribute non-financial information, such as annual reports, presentations to financial analysts or corporate websites.

Several economic theories provide explanations for differences in the amount and in the complexity of voluntary non-financial disclosures disseminated by firms (Khlifi and Bouri, 2010). Agency theory contends that firms are more likely to be transparent when agency conflicts between insiders and outsiders are larger since these conflicts lead to higher levels of information asymmetry. Empirical studies hence confirm that the voluntary disclosure of non-financial information has a positive association with the dispersion in firms' ownership structures (Marston and Polei, 2004; Prencipe, 2004) and with firm size (García-Meca et al., 2005). Signalling theory arguments that managers of firms with higher financial performance disclose more information voluntarily in order to promote a positive image (Mohd Ghazali and Weetman, 2006). Voluntary disclosure provides good signals about future firm

performance and avoids the risk that outsiders make wrong judgments based on non-disclosure of corporate information (Khlifi and Bouri, 2010). Empirical evidence confirms a positive association between firm performance and the reporting of non-financial information (Soliman, 2013). Legitimacy theory arguments a higher voluntary publication of non-financial information, and especially related to CSR activities, to legitimize firms' activities and to respond to social pressures (Brown and Deegan, 1998). Empirical results document higher levels of environmental disclosure from firms with environmentally sensitive production activities (Aerts et al., 2008; Brammer and Pavelin, 2008). Finally, the proprietary cost theory contends that firms are reluctant to convey corporate non-financial information voluntarily when competitive costs are larger which are detrimental to firms.

The selective disclosures of non-financial information among firms arises the question whether such information should be regulated to a larger extent (Cohen et al., 2012). Despite the increased relevance of non-financial information in firm valuation, ample legislative initiatives are undertaken to require firms to report non-financial information. To illustrate, the current European Union (EU) legislation, only provides one article, i.e. article 46(1)(b) of the fourth Directive stating that: "To the extent necessary for an understanding of the company's development, performance or position, the analysis [in the annual review] shall include both financial and, where appropriate, non-financial key performance indicators relevant to the particular business, including information relating to environmental and employee matters". Based on this article, the legal framework at EU Member States' level appears to be quite fragmented. A few Members States have adopted mandatory reporting obligations, with different approaches varying from very detailed reporting requirements to more general provisions (EC, 2011b). Some Member States have introduced disclosure requirements that go beyond the Fourth Directive. Other Member States have made the disclosure of non-financial information mandatory. Still others have adopted a 'comply or explain' regime. The Member States also have the opportunity to exempt small and medium-sized firms from this requirement.

In 2011, the EU took the initiative to regulate the disclosure of social and environmental information so as to improve the comparability, the reliability and the usefulness of non-financial information (EC, 2011a). In April 2014, the European Commission adopted a directive on the disclosure of non-financial information by certain large firms. The commission in particular stimulates EU large firms to disclose information related to environmental aspects, social and employee related topics, respect for human rights, anti-corruption and bribery issues, and diversity on board of directors in the annual reports (EC, 2014)

A potential advantage of endorsing non-financial information reporting requirements is to enhance the reliability of the information reported, if such information is also assured by professional accountants (Cohen et al., 2012). Reporting standards would bring consistency to reporting and permit comparability of information, at least within industries. In addition, a standard would provide a benchmark against which reports could be assessed and assurance could be provided. However, although the lack of a generally accepted framework to report non-financial information is an important barrier to widespread acceptance and use of non-

financial information by investors and other stakeholders, such information is hard to mandate and to standardize. Despite the reporting requirements, firms could limit their efforts by publishing vague and uninformative disclosures. In addition, a common framework is irrelevant since the importance of non-financial information depends on firm and industry characteristics. In addition, non-financial information requirements increase accompanied costs such as reporting or proprietary costs (Skinner, 2008; Stark, 2008). So, notwithstanding some advantages, non-financial information reporting requirements are quite limited, and even when requirements exist, these are not strictly enforced (Moser and Martin, 2012).

Instead of reporting requirements to stimulate firms to report non-financial information, various initiatives have recommended firms to disclose non-financial information voluntarily. Voluntary non-financial disclosure is considered to be more effective in improving the efficient functioning of capital markets rather than mandating non-financial disclosure (Bushee and Leuz, 2005; Ahmed and Schneible, 2007; Gomes et al., 2007; Skinner, 2008). During the past two decades, many ideas for improving business reporting have been issued and nearly all of them focus on releasing more non-financial information. Since 2008, at least 18 organizations have issued frameworks and guidance for reporting non-financial information (Eccles et al., 2011). Diverse authorities, such as the American Institute of Chartered Accountants (AICPA), the Financial Accounting Standards Board (FASB) or the Institute of Chartered Accountants in England and Wales (ICAEW) developed reporting models in which recommendations were made for the provision of non-financial indicators. Other recommendations concerning the voluntary reporting of non-financial information, are, for instance, the Global Reporting Initiative (GRI), the UN Global Compact, the OECD Guidelines for multinational firms, the ISO 26000 Guidelines, the UN 'Protect, Respect or the Remedy' Framework for Business and Human Rights (commonly referred to as 'Ruggie Framework'). In response to the call from organisational stakeholders to submit non-financial information, an increasing number of firms have been experimenting with more robust disclosure of non-financial information. The GRI Sustainability Reporting Guidelines, better known as G3, may be the most widely used framework to convey non-financial information (Eccles et al., 2011). G3 provides guidance on reporting on firms' economic, environmental, and social performance.

Although it seems that a certain flexibility may need to be maintained, flexibility is not optimal either. The proliferation of guidance raises another issue. This existence of different frameworks creates a perception about 'competing frameworks' and causes confusion in the marketplace about what framework a company should use.

To sum up, extant literature highlights the increasing importance and usefulness of non-financial information in the decision-making process of various stakeholders over time (Cohen et al., 2012). Non-financial information complements the financial information stakeholders have at their disposal. For instance, investors rely upon non-financial information to judge firms' future cash flows and value creation potential or to assess the board competencies (Cohen et al., 2011; Fernández et al., 2011). Customers and employees rely on corporate social performance in order to make purchase decisions (Schuler and Cording, 2006), to apply for a

job position (Turban and Greening, 1996) or to increase employee commitment (Dogl and Holtbrugge, 2014). Firms hence realize that they need to respond to the expectations and concerns of the various stakeholders with whom they interact, not only by adapting their activities, but also by reporting how they have performed. As such, they will respond to the question of stakeholders to publish all kinds of non-financial information on a voluntary basis.

Economic theory also contends that firms could profit from improved transparency about their non-financial value drivers in terms of higher performance and value. A greater transparency leads to an improved confidence, image and reputation by organizational stakeholders such as investors, employees and customers leading to better firm performance and firm value (Cormier et al., 2009; Orens et al., 2010; Dhaliwal et al., 2012; Athanasakou and Hussainey, 2014). The public reporting of non-financial information however attracts costs, such as reporting costs and proprietary costs. Hence, it is useful to firms to comprehend the extent to which non-financial information is used by their stakeholders. Since non-financial information is a broad concept, corporations and regulators are interested in the types of non-financial information stakeholders find useful in their decision-making process. In order to allow firms to judge whether the non-financial information conveyed is useful, the remainder of the literature review focuses on studies examining the use of corporate non-financial information by financial analysts. Understanding their preferences helps to improve the future information flow between firms and their stakeholders (Barker, 1998).

III. THE USE OF NON-FINANCIAL INFORMATION BY FINANCIAL ANALYSTS

This section of the literature review details prior findings on the extent to which financial analysts rely upon non-financial information in making firm assessments. Financial analysts are primary users of corporate information and are considered as the representatives of the investment community for whom the reporting of corporate information is primarily intended (Schipper, 1991; Chen et al., 2010).

Financial analysts add value to investors by disseminating information about the firm and by monitoring firm management (Livnat and Zhang, 2012). In serving both functions, financial analysts affect the investors' decision-making process (Hirst et al., 1995; Ackert et al., 1996; Holland and Johanson, 2003; Ivković and Jegadeesh, 2004; Covrig and Low, 2005; Fogarty and Rogers, 2005). Financial analysts alleviate the information asymmetry between investors and firms, and add additional knowledge to the information that is publicly disclosed by firms (Barber et al., 2001; Rammath et al., 2008), increasing the efficient functioning of capital markets (Barker, 1998; Holland and Johanson, 2003).

The following sections elaborate on studies investigating the analysts' use of non-financial information conveyed by firms. First, literature findings are reported which indirectly examined this use by relating corporate non-financial disclosures to properties of analysts'

earnings forecasts such as their accuracy and dispersion. Significant associations between the reporting of non-financial information and these properties indicate that financial analysts rely on non-financial information to predict future earnings, and hence consider such information relevant in their decision making process. The next section reports empirical studies addressing the financial analysts' use of corporate non-financial information directly through questionnaires and the content analysis of analyst reports.

A. DISCLOSURE OF NON-FINANCIAL INFORMATION AND ITS INFLUENCE ON THE PROPERTIES ON ANALYSTS' EARNINGS FORECASTS

This section reports studies examining the association between the voluntary disclosure of non-financial information and the properties of analysts' earnings forecasts. Economic theory suggests a reduction in the uncertainty financial analysts perceive about firms' future earnings or cash flows with increased discretionary disclosure (Barry and Brown, 1985; Leuz and Verrecchia, 2000; Bushman and Smith, 2001). Consistent with the view that improved disclosure is inversely associated with the level of information asymmetry (Barry and Brown, 1985; Leuz and Verrecchia, 2000), greater disclosure allows financial analysts to increase their knowledge about the firms' activities (Hope et al., 2006). As far as the information conveyed by firms is relevant, financial analysts include this information into their valuation models to judge future firm performance (Cormier and Magnan, 2013). The more relevant information is reported by firms, the less uncertainty financial analysts experience in forecasting firms' future earnings and value (Bushman and Smith, 2001; Hope et al., 2006), resulting in more accurate and less dispersed earnings forecasts. In addition, firms improving their disclosure strategy attract a larger number of financial analysts increasing the competition between these analysts to serve investors' needs. This motivates financial analysts to increase their efforts in collecting corporate information, which further reduces the uncertainty about future firm performance of firms, leading to an improvement in the forecast accuracy and a decline in the forecast dispersion. So in general a positive (negative) association between the extent of non-financial information disclosure and the accuracy (dispersion) of the forecasted earnings is assumed. However, Barron et al. (2002) also pointed out that improvements in the level of public disclosures encourage analysts to collect private information which could lower the consensus across financial analysts, translating into an increase in the forecast dispersion. In addition, if financial analysts judge non-financial information differently, forecast dispersion could also increase as well. So from a theoretical point of view, the negative association between the extent of voluntary disclosure and forecast dispersion is unclear. To judge the expected associations, Table 1 summarizes empirical research findings related to the link between the voluntary disclosure of non-financial information and the properties of analysts' earnings forecasts.

Table 1: Disclosure of non-financial information and properties of analysts' earnings forecasts

Authors	Year	Sample	Type of non-financial information	Key results
Lang & Lundholm	1996	USA	– Financial and product related information in annual reports – Financial information released in other publications (quarterly reports, press releases, presentations to analysts) – Investor relations information	– Negative association between forecast dispersion and annual report information and investor relations information. – Positive association between forecast accuracy and the content of other publications and investor relations.
Barron, Kile & O'Keefe	1999	USA	MD&A quality (general financial and non-financial information) in annual reports	– MD&A ratings are positively associated with forecast accuracy and negatively associated with forecast dispersion
Vanstraelen, Zarzeski, & Robb	2003	Continental Europe (Belgium, Germany, the Netherlands)	– Forward-looking non-financial information (environment, strategy, management, trends) in annual reports – Historical non-financial information (business structure, production, customers) in annual reports	– Negative (positive) association between forward-looking nonfinancial information and forecast dispersion (accuracy). – Historical non-financial information is unrelated with forecast dispersion and forecast accuracy.
Aerts, Cormier & Magnan	2007	North-America (USA and Canada) Continental Europe (Belgium, France, Germany, the Netherlands)	Web-based performance disclosure index including indicators related to: – Financial performance – Corporate governance – Customer value – Human/intellectual capital – Production efficiency, innovation & R&D – Social responsibility	– Negative association between performance disclosure index and forecast dispersion, but only for firms from North America. – Insignificant association between performance disclosure index and forecast dispersion in Continental Europe. – Negative association between performance disclosure index and forecast dispersion is lower in firms with high analyst following.

Authors	Year	Sample	Type of non-financial information	Key results
Aerts, Cormier & Magnan	2008	North-America (USA and Canada) Continental Europe (Belgium, France, Germany, the Netherlands)	Environmental reporting in annual reports and corporate websites	– Negative association between environmental disclosure and forecast dispersion in North America and Continental Europe. – Negative association is attenuated by increased analyst following. – Negative association in North America is only observed for print based environmental disclosure and not web-based disclosure. – Positive association between environmental disclosure and forecast dispersion in firms from Continental Europe operating in environmentally more sensitive industries.
Bozzolan, Trombetta & Beretta	2009	Continental Europe (Italy, Germany, France and Switzerland)	Forward-looking information in annual reports	– Verifiable (quantified) disclosure improves forecast accuracy and forecast dispersion more in comparison to unverifiable disclosure.
Nichols & Wieland	2009	USA	Product-related information (product announcements such as new products, new developments, product discontinuation or product recalls) and business expansion information in press releases	– Negative (positive) association between product-related disclosure and business expansion information and forecast dispersion (accuracy).
Luo, Homburg & Wieseke	2010	USA	Customer satisfaction	– Negative association between customer satisfaction and forecast dispersion. – Negative association is stronger with more competition and larger market uncertainty.
Simpson	2010	USA	Customer related performance metrics (i.e. number of subscribers, market share, customer acquisition cost, average revenue per user, churn rate, and minutes of use per subscriber)	– Positive association between customer acquisition cost, average revenue per user and number of subscribers and analysts' earnings forecasts. – Other metrics are unrelated to analysts' forecast accuracy.

Authors	Year	Sample	Type of non-financial information	Key results
				– The customer related performance metrics are negatively related to analysts' forecast dispersion
Yu	2010	Global	Corporate governance information in annual reports	– Negative (positive) association between corporate governance disclosure and forecast dispersion (accuracy)
Dhaliwal, Radhakrishnan, Tsang & Yang	2011	USA	Issuance of a stand-alone corporate social responsibility (CSR) report	– Negative (positive) association between the issuance of a CSR report and forecast dispersion (accuracy)
Hsu & Chang	2011	Taiwan (high tech firms)	– Intellectual capital information in annual reports – Non-intellectual capital disclosure (general information, social responsibility, management discussion and analysis and historical results) inserted in annual reports	– Negative (positive) association between intellectual capital disclosure and forecast dispersion (accuracy) – Insignificant association between non-intellectual capital disclosure and properties of analysts' earnings forecasts
Dhaliwal, Radhakrishnan, Tsang & Yang	2012	Global	Issuance of a stand-alone corporate social responsibility (CSR) report	– Positive association between issuance of a CSR report and forecast accuracy – Association between CSR report and forecast accuracy is stronger in countries with stakeholder orientation (i.e. countries where CSR performance affects firm performance) – Association between CSR report and forecast accuracy is stronger in countries with more opaque financial disclosure
Ngobo, Casta & Ramond	2012	USA	Customer satisfaction	– Positive association between customer satisfaction and forecast accuracy.
Cormier & Magnan	2013	Canada and USA	Environmental information in annual and sustainability reports	– Positive association with the consensus of earnings forecasts (less forecast dispersion). – Consensus is reduced when information is disclosed by firms with poor environmental performance.

The findings revealed in Table 1 tend to confirm the proposition that forecast accuracy improves and forecast dispersion lowers with expanded non-financial information disclosures. In a US setting, Lang and Lundholm (1996) observe that Financial Analysts Federation (FAF) disclosure ratings have a negative association with the errors and the dispersion in the earnings forecasts. Barron et al. (1999) show that the analysts' forecast dispersion is decreasing and analysts' forecast accuracy is increasing with better disclosures in the Management Discussion and Analysis (MD&A).

However, institutional differences affect the association between the extent of non-financial information reported and the financial analysts' earnings forecast dispersion. Whilst a negative association is observed between the disclosure of non-financial information on corporate websites and the dispersion in the analysts' earnings forecasts in North America, an insignificant association is shown in Continental Europe (Aerts et al., 2007). The latter result confirms the assumption that (1) common law institutional regimes are more focused on full transparency to inform investors compared to code law countries which are more characterized to formal compliance with existing requirements and hence low levels of disclosure (Basu et al., 1998) and (2) the analysts' labour market in Continental Europe provides fewer incentives to predict earnings forecasts accurately compared to North America (Clement et al., 2003; Bolliger, 2004). In addition, the association between web-based non-financial disclosure and the level of forecast dispersion is attenuated with analyst following, indicating that analysts following acts as a substitute for corporate disclosures (Hope, 2003). In contrast with previous results, a negative association between financial analysts' forecast dispersion and the extent of non-financial disclosure is observed for a sample of Belgian firms (Orens and Lybaert, 2010a).

Since non-financial information is a broad concept, we focus in this literature review to the usefulness of some specific categories of non-financial information to value firms and to forecast earnings. In particular, we focus on following types of non-financial information: strategic and product related information, forward-looking information, corporate governance information, intellectual capital information, and social and environmental information. These non-financial information categories are often researched in the voluntary disclosure literature. Regarding the strategy-related and product-related information, Nichols and Wieland (2009) find for a sample of US firms that the reporting of product related information and business expansion information in press releases allow financial analysts to make more accurate and less dispersed earnings forecasts. These disclosures ameliorate analysts' impressions about future sales and earnings. In a code law country setting, disclosure of general business information (for instance about firms' products or markets) tends to be unrelated with the properties of analysts' earnings forecasts (Orens and Lybaert, 2010a).

Empirical evidence further reveals that greater forward disclosure leads to less dispersed and more accurate earnings estimates for a sample of Continental European firms (Vanstraelen et al., 2003; Bozzolan et al., 2009; Orens and Lybaert, 2010a). Additionally, verifiable (or quantifiable)

forward-looking disclosure shows a stronger association with the accuracy and dispersion of the analysts' earnings forecasts compared to unverifiable (or descriptive) information. The results tend to indicate that quantifiable forward-looking disclosures, which could be compared with actual results, are more relevant to financial analysts (Bozzolan et al., 2009)

Yu (2010) observes that financial analysts predict more accurate and less dispersed earnings forecasts when firms disclose more comprehensive corporate governance disclosures (for instance about their board composition, ownership structures or managerial characteristics). Analysts tend to rely upon corporate governance information in estimating future earnings. Corporate governance disclosures allow analysts to assess the firms' board policy and the firms' potential risks and future prospects (Durnev and Kim, 2005). More knowledge about corporate governance reduces analysts' uncertainty about the firms' future prospects. Yu (2010) comments that the research results are dominated by US firms driving the research results, but a negative association between corporate governance disclosures and forecast dispersion is also observed for a sample of Continental European firms (Orens and Lybaert, 2010a).

Focusing on intellectual capital information, Hsu and Chang (2011) observe a lower diversity of beliefs across financial analysts when firms from high tech industries submit more intellectual capital information. In addition, financial analysts forecast more accurately when firms publish more intellectual capital disclosures. Aerts et al. (2007) document a negative relationship between the disclosure of intellectual capital information and the analysts' forecast dispersion, but only for firms from North America and not for Continental European firms. However, this result is inconsistent with Orens and Lybaert (2010a) documenting that the disclosure of intellectual capital information is associated with less dispersed analysts' forecasts for a sample of Belgian firms.

Luo et al. (2010) find that firms with better customer satisfaction scores, provided by the American Customer Satisfaction Index (ACSI), exhibit more consensus across the earnings forecasts prepared by financial analysts. Positive changes in customer satisfaction are expected to increase future firm performance – since future cash flows are more vulnerable – leading to improved decision making by financial analysts. In addition, this association is stronger in more competitive industries since high levels of customer satisfaction in these industries enhance sales levels relative to their competitors. In less competitive industries, it is more likely that customer relationships retain, despite low levels of customer satisfaction. Hence, customer satisfaction data is more important in industries with high product market competition. Ngobo et al. (2012) confirm findings in Luo et al. (2010), showing a positive association between customer satisfaction and forecast accuracy.

To understand the influence of corporate social responsibility (CSR) information on the predictions of financial analysts, Dhaliwal et al. (2012) investigate the impact of the issuance of a stand-alone CSR report on the accuracy of the earnings forecasts. Using a sample of firms from 31 countries from all continents, they observe that financial analysts predict earnings more accurately when firms publish a CSR report. Since CSR performance impacts future results, information related to CSR activities is useful, leading to more informed earnings estimates.

The results confirm earlier findings from Dhaliwal et al. (2011) observing a negative association between CSR disclosure on the one hand and the extent of forecast errors and forecast dispersion on the other hand for a sample of US firms. In addition, Dhaliwal et al. (2012) observe a stronger association between the disclosure of a separate CSR report and the accuracy of the analysts' earnings forecasts in countries that are more stakeholder oriented compared to countries that are more shareholder-oriented. In a stakeholder-oriented environment, stakeholders have a greater influence on the activities of firms in comparison to shareholder-oriented countries (Chen, 2009). In addition, Dhaliwal et al. (2012) note that the disclosure of CSR is complementary to financial disclosure since the relationship between the disclosure of a CSR report and the forecast accuracy is stronger in countries with a higher level of financial opacity. CSR information hence mitigates the negative impact of financial opacity on forecast accuracy.

Aerts et al. (2008) further find a higher consensus across financial analysts' earnings forecasts if firms from both Continental Europe and North America disclose more environmental information, but this association is attenuated when analyst following is larger. Comparable results are observed in Cormier and Magnan (2013) finding that the consensus among financial analysts is larger when firms convey more environmental disclosures. These results are obtained in a North American setting.

To sum up, extensive non-financial information reporting enriches the information environment, improves the predictability of future earnings and reduces asymmetry across analysts in their beliefs about future prospects. An increase in the voluntary disclosure of non-financial information leads to a decline in the earnings forecast dispersion and a decline in the earnings forecast errors. Hence, firm management possibly could profit from a lower cost of capital since the reduction in the forecast dispersion results in lower uncertainty levels across investors (Khurana and Raman, 2004; Gietzman and Ireland, 2005). Extant literature also shows that the disclosure of non-financial information has a stronger impact on the consensus of the earnings forecasts than on the forecast accuracy of the earnings estimates. This finding is due to the reduction in the private information flow between firms and a selected number of financial analysts, resulting in more consensus across analysts (Barron et al., 1999). Potentially, the diverse interpretation of non-financial information by financial analysts does not immediately infer more accurate earnings estimates.

B. USE OF CORPORATE NON-FINANCIAL INFORMATION BY FINANCIAL ANALYSTS: SURVEY AND CONTENT ANALYSIS EVIDENCE

To observe financial analysts' use of non-financial information directly, some studies content analyse the reports issued by financial analysts. Studies making use of this approach assume that the information elements discussed in these reports reflect the most important ones financial analysts use in assessing firms (Rogers and Grant, 1997; Bradshaw, 2004;

Abhayawansa, 2011). An analyst report[3] should include all the necessary information a financial analyst uses to argument its stock recommendation (García-Meca, 2005). Based on a detailed analysis and interpretation of the collected corporate information, financial analysts issue an analyst report including a description of a firm's profile, an opinion on the current and future firm performance, an estimation of the firm's future earnings and cash flows and a stock recommendation (Asquith et al., 2005).

The content analysis approach might create some bias since no conclusions can be drawn as to whether financial analysts include all information they use (or consider useful) in their reports (Rogers and Grant, 1997; Abhayawansa, 2011). To validate the content analysis results, the survey approach is used. One drawback of this research method might be that the research findings do not correspond with financial analysts' actual behaviour. However, empirical results document a strong correlation in the use of corporate non-financial information collected based on a survey and based on a content analysis of reports prepared by the survey respondents. Non-financial information elements being used more frequently according to the survey, are more frequently inserted in the analyst reports (Orens and Lybaert, 2007). Table 2 synthesizes prior empirical results about the use of corporate non-financial information using either the content analysis approach or the survey approach.

In general, results tend to indicate that financial analysts increasingly use corporate non-financial information over time. Studies conducted in the nineties reveal a very limited use of non-financial information by financial analysts. In that period, analyst reports only include product-related information, market-related information and forward-looking information about the opportunities and risks (Previts et al., 1994; Rogers and Grant, 1997; Breton and Taffler, 2001). More recent studies document an evolution in the use of non-financial information, and observe that a substantial part of an analyst report is attributed to a discussion of non-financial information (García-Meca, 2005; Flöstrand, 2006; García-Meca and Martinez, 2007; Orens and Lybaert, 2007; Abhaywansa and Guthrie, 2012).

Considering the different types of non-financial information, analyst reports often report product related information (García-Meca, 2005; Orens and Lybaert, 2007; Nielsen, 2008). A strong variability exists in the extent to which analyst reports disseminate detailed information about firms' strategy (Flöstrand and Ström, 2006; Orens and Lybaert, 2007). Although survey evidence confirms the importance of strategy-related information for financial analysts to value a firm, proprietary costs hamper firms to report such information publicly, which in turn limits the dissemination of strategy related information in analyst reports (Dempsey et al., 1997; Breton and Taffler, 2001; Sakakibara et al., 2010).

[3] In general, two types of analyst reports exist: company reports and result reports (García-Meca and Martinez, 2007). Company reports include much more corporate information compared to result reports since financial analysts in these reports present a fundamental analysis of firms providing a detailed picture of firms' activities and performance. Financial analysts however do not publish such reports on a regular basis. Result reports are published more frequently during the year and include information related to a particular event, for instance an earnings announcement, the launch of a new product or an acquisition.

Table 2: Use of non-financial information by financial analysts using direct research methods

Authors	Year	Sample	Method	Key results
Previts, Bricker, Robinson & Young	1994	USA	Content analysis (479 analyst reports)	– Few attention to non-financial information, only focus to forward-looking information, competitive position, strategy and management
Rogers & Grant	1997	USA	Content analysis (187 analyst reports)	– Large use of background information (products, markets, stakeholders) and forward-looking information – Less attention to corporate governance information and managers' analysis of firm performance
Dempsey et al.	1997	USA	Survey (240 respondents)	– Frequent use of information relating to the competitive environment, the quality of firm management and product innovation – Information related to human capital, social responsibility and customer satisfaction is hardly used.
Breton & Taffler	2001	UK	Content analysis (105 analyst reports)	– Frequent use of product market information – Strategy and management most important information, but to a lower extent reported in analyst reports
García-Meca	2005	Spain	Content analysis (217 analyst reports)	– Analyst rely most on strategy and objectives – Analysts frequently use product-related information and customer value information – Analysts rarely use human capital information
Flöstrand	2006	USA	Content analysis (250 analyst reports)	– Customer value information most employed – Few attention to human capital and internal structure (innovation, R&D) information
Flöstrand and Ström	2006	USA	Content analysis (200 analyst reports)	– Forward-looking information and background information most used – Strategy information is hardly included
García-Meca and Martínez	2007	Spain	Content analysis (217 analyst reports)	– Strategy and forward-looking information is often used – Innovation, R&D and human capital receive less attention

Authors	Year	Sample	Method	Key results
Orens and Lybaert	2007	Belgium	Content analysis (177 analyst reports) and survey (31 responses)	– Forward-looking information and background information (strategy and product information) often used – Management's analysis of firm performance frequently employed – Corporate governance information and intellectual capital information hardly used
Nielsen	2008	Denmark	Content analysis (12 analyst reports)	– Large use of general background information (products, markets, industries) – Management's analysis of firm performance and forward-looking information frequently used – Few attention addressed to intellectual capital information, corporate governance and social responsibility information
Abhayawansa & Guthrie	2012	Australia	Content analysis (62 analyst reports)	– External and human capital employed frequently – Internal capital hardly disclosed – Use of intellectual capital information conditioned by stock recommendation (e.g., unfavourable stock recommendations include more external capital information)
Pinho, Madaleno & Santos	2013	Portugal	Content analysis (73 analyst reports)	– Analyst rely heavily on management's discussion of firm performance and product related information – Intellectual capital information, corporate governance information, and information about risks and opportunities are rarely conveyed in analyst reports
Abhayawansa & Guthrie	2014	Australia	Content analysis (64 analyst reports)	– Information on relational capital and company management commonly used – Information on human capital and structural capital used least frequently

The empirical studies further notice a strong reliance on forward-looking information in analyst reports (Orens and Lybaert, 2007; García-Meca and Martinez, 2007; Nielsen, 2008). Forward-looking information is a guide to assess the long term performance of firms. Questionnaire results confirm the importance of forward-looking information in firm assessments (Orens and Lybaert, 2007). Despite the importance of corporate governance on firm performance, analysts do not mention such information in their analyst reports (Orens and Lybaert, 2007; Nielsen, 2008). Since such information is often disclosed by public firms, financial analysts could be reluctant to communicate such information through the analyst reports. Despite the limited occurrence of corporate governance information in the analyst reports, survey evidence tends to confirm that financial analysts rely upon information about top management in assessing firms (Dempsey et al., 1997; Orens and Lybaert, 2007).

Focusing on intellectual capital information, Nielsen (2008) observes an infrequent use of this information in analyst reports. However, intellectual capital could be broken down into three groups: human capital (for instance, employee satisfaction, training), internal (or structural) capital (for instance innovation or R&D) and external (or relational) capital (for instance customer value information). In general, extant literature shows an infrequent use of human capital and internal capital information (García-Meca, 2005; Flöstrand and Ström, 2006; Flöstrand, 2006; García-Meca and Martínez, 2007; Orens and Lybaert, 2007; Abhayawansa and Guthrie, 2014). Abhayawansa and Guthrie (2012) is the only study documenting a frequent use of human capital information in the reports analysed. When intellectual capital information is disclosed, most attention is addressed to external capital information (Breton and Taffler, 2001; García-Meca, 2005; Flöstrand and Ström, 2006; Flöstrand, 2006; Orens and Lybaert, 2007; Abhayawansa and Guthrie, 2012).

Although prior studies confirm the usefulness of intellectual capital information in assessing firm value (for instance, Barth and Clinch, 1998; Kallapur and Kwan, 2004), financial analysts have their reservations about the validity and reliability of intellectual capital information, which makes it difficult to gauge the impact of this information on future cash flows (Barker, 1998; Johanson, 2003). In addition, due to proprietary costs, firms are less likely to disclose intellectual capital information, such as customer satisfaction or product quality, publicly (García-Meca, 2005), increasing collection costs for financial analysts (Dempsey et al., 1997, Orens and Lybaert 2007). In order to gain insights into the impact of intellectual capital information on firm value, financial analysts have to collect this information privately. Finally, lack of knowledge and experience in assessing intellectual capital information and its link with firm value, might also explain why intellectual capital information is hardly employed (Holland, 2003).

Corporate social responsibility information is hardly included in an analyst reports (Nielsen, 2008). Based on a survey, Hunt and Grinnel (2003) confirm these findings with respect to environmental information. A potential explanation for the limited use of this information relates to the low credibility of this information (Hunt and Grinnel, 2003).

Conducting interviews, Campbell and Slack (2011) observe that financial analysts ignore environmental information when evaluating firms.

To summarize, financial analysts often use product-related information and forward-looking information (e.g. about future products or future opportunities and risks). Information about the strategy and objectives of firms is largely considered as well by financial analysts in their firm evaluation, but proprietary costs reduce the availability of this information for financial analysts to insert the information in their analyst reports. Previous literature further shows a low attention addressed to corporate governance information, intellectual capital information and corporate responsibility information in analyst reports. Potentially, financial analysts have a short term orientation and neglect information that consider long-term issues, such as environmental information, to judge firms (Campbell and Slack, 2011; Arvidsson, 2012).

The increased dissemination of non-financial information in analyst reports is possibly due to the regulatory changes regarding the working environment of financial analysts. Following the scandals in the 1990s, some rules were endorsed to restrict contact between financial analysts and bankers and to strengthen the "Chinese wall" separating equity research and investment banking (Brown et al., 2014). The avoidance of conflicts of interests allows financial analysts to be more critical towards firms, potentially leading to more developed equity reports. In addition, professional organisations, such as the Certified Financial Analysts (CFA) Institute, push financial analysts to issue objective reports and recommendations that are supported by a thorough research and investigation. Following the various guidelines and regulations, financial analysts are under more pressure to disseminate objective analyst reports. As the relevance of financial information in equity valuation reduces, in favour of non-financial information, financial analysts might be more keen to report non-financial information in their analyst reports.

To understand the context within which financial analysts make decisions about the levels of non-financial information used, several studies address potential clarifications for differences in the financial analysts' use of corporate non-financial information. Flöstrand (2006) shows that analyst reports issued for firms in the pharmaceutical industry and the telecommunications industry contain more intellectual capital information compared with analyst reports on energy firms. Industry membership hence affects the relative importance of non-financial information. García-Meca and Martinez (2007) find that the amount of non-financial information in the analyst reports is increasing with firms' profitability and growth opportunities. García-Meca and Martínez (2007) conclude that financial analysts release more intellectual capital information in their reports when a sell recommendation is at stake compared to a buy recommendation. Abhayawansa and Guthrie (2012) build further on these insights relating the theory of impression management to investigate how type and level of detail of intellectual capital information vary by recommendation type. Analysts employ intellectual capital information in their reports to manage perceptions. So, analysts communicate more external capital information in their reports with an unfavourable stock recommendation. Favourable recommendations include more future oriented intellectual

capital information compared to unfavourable recommendations containing more historical information (Abhayawansa and Guthrie, 2012). Orens and Lybaert (2010b) demonstrate that financial analysts use more non-financial information when doubts arise about the reliability of the earnings figures. In addition, less experienced financial analysts and analysts covering fewer firms rely upon a larger amount of corporate non-financial information.

IV. CONCLUSION

The current literature review focuses on the use of corporate non-financial information by financial analysts. Non-financial information is increasingly important due to the transition towards a knowledge economy, the globalization or the intensified competition. Various corporate stakeholders call for the voluntary disclosure of non-financial information since the financial information included in financial statements is insufficient to judge future firm performance and firm value accurately. Non-financial information should hence complement the financial information that is reported on a mandatory basis.

To understand the relevance of corporate non-financial information in the decision-making process of financial analysts, some studies examine the influence of the voluntary disclosure of non-financial information on the properties of analysts' earnings forecasts, such as the accuracy and the dispersion of these forecasts. Another approach to discover the reliance of financial analysts on non-financial information is to directly analyse the output of financial analysts, i.e. the analyst reports, or to submit questionnaires to financial analysts. Both approaches suggest the usefulness of non-financial information to financial analysts in analysing the current and future performance of firms. Findings reveal a decline in the forecast dispersion and an increase in the forecast accuracy with improved non-financial disclosures. However, this association tends to be more pronounced in a common law setting.

Considering the various types of non-financial information, financial analysts mainly use forward-looking information and strategic and product related information. Financial analysts to a minor extent rely upon intellectual capital information, corporate governance information and social and environmental information. Hence, it seems that analysts fail to include or consider these information elements in their evaluations although these indicators are considered important in firm valuation.

The findings provide important implications for firm management making decisions about the disclosure policy to be followed. The results tend to indicate that the relevance of corporate non-financial information differs across types of non-financial information. Disseminating useful information to capital markets could be favourable to firms in terms of lower information asymmetry, lower cost of capital, higher financial performance and higher firm value.

Analysing prior studies exhibits the finding that non-financial information remains descriptive, which hampers the use of this information by capital market participants. Quantification of non-financial information is recommended since such information

increases the credibility of the information provided as it improves the ex post verifiability of the information disclosed (Hutton et al., 2003). Another option to facilitate the use of corporate non-financial information is that firms should take efforts to better clarify the relationship between non-financial information and firm value (Dickins and Higgs, 2005).

Reflecting on topics for further research, we suggest to concentrate more on the association between properties of analysts' earnings forecasts and non-financial disclosure in other communication venues such as conference calls, press releases, management forecasts or webcasts. Prior studies mainly focus on the disclosure of non-financial information in annual reports to obtain insights into the extent of non-financial information reporting. Despite previous results, it is still a black box how non-financial information is considered as input into forecasts and stock recommendations by financial analysts. Future research could for instance adopt a protocol analysis allowing to capture detailed insights into how analysts incorporate non-financial information in their forecasts and which tools are used to value the implications of non-financial information. In addition, there is ample evidence about the information sources upon which financial analysts rely to gather non-financial information. It is important to gain more insight into the extent to which financial analysts collect non-financial information privately or publicly. Such research maps the potential information asymmetry between investors and financial analysts.

Most studies use a pre-defined list of non-financial information items and identify the extent to which each information element is used by financial analysts. However, the construction of this list might be constrained by the researchers' judgement to define and categorize non-financial information (Abhayawansa, 2011). An alternative approach to content analyze analyst reports is to consider all non-financial information elements included in the analyst reports, and group these elements into various non-financial information categories. In addition, it is still unclear for which purposes each financial analyst relies on a non-financial information element. For instance, financial analysts might include a non-financial information item in their reports either with the intention to provide some background information about the firm to investors or with the intention to use it as input for firm valuation (Abhayawansa and Guthrie, 2012).

Finally, the question arises whether additional information requirements should be set. Recently the European Commission launched a directive to require firms to disclose social and environmental related information, but still many non-financial information elements such as forward-looking information or product related information have to be reported voluntarily. However, regulators face difficulties in setting non-financial information requirements as the importance of non-financial information is depending on firm and industry characteristics (Skinner, 2008; Stark, 2008). A common framework including non-financial information would be irrelevant for all firms (Stark, 2008). The literature review also provides evidence that the emphasis placed on non-financial information by financial analysts is conditioned by the nature of the covered firms. In other words, firm-specific factors drive the decision of financial analysts to use non-financial information. This finding allows us to suggest that setting information requirements for all firms is ineffective.

REFERENCES

Abhayawansa, S., 2011. A methodology for investigating intellectual capital information in analyst reports. Journal of Intellectual Capital, 12(3), 446–476.

Abhayawansa, S., Guthrie, J., 2012. Intellectual capital information and stock recommendations: impression management? Journal of Intellectual Capital, 13(3), 398–415.

Abhayawansa, S., Guthrie, J., 2014. Importance of intellectual capital Information: A study of Australian analyst reports. Australian Accounting Review, 24(1), 66–83.

Ackert, L., Church, B., Shehata, M., 1996. What affects individuals' decisions to acquire forecasted information? Contemporary Accounting Research, 13 (2), 379–399.

Adams, A., Fries, S., Simnett, R., 2011. The journey towards integrative reporting. Accountant's Digest, 558, 1–41.

Aerts, W., Cormier, D., Magnan, M., 2007. The association between web-based corporate performance disclosure and financial analyst behavior under different governance regimes. Corporate Governance: An International Review, 15(6),1301–1329.

Aerts, W., Cormier, D., Magnan, M., 2008. Corporate environmental disclosure, financial markets and the media: An international perspective. Ecological Economics, 64(3), 643–659.

Ahmed, A., Schneible, R., 2007. The impact of regulation fair disclosure on investors' prior information quality – evidence from an analysis of changes in trading volume and stock price reactions to earnings announcements. Journal of Corporate Finance, 13 (2–3), 282–299.

American Institute for Certified Public Accountants (AICPA), 1994. Improving business reporting: a customer focus: Meeting the information needs of investors and creditors – Report of the Special Committee on Financial Reporting. AICPA, New York.

Amir, E., Lev, B., 1996. Value-relevance of nonfinancial information: The wireless communications industry. Journal of Accounting and Economics, 22(1), 3–30.

Arvidsson, S., 2012. The corporate communication process between listed companies and financial analysts. Corporate Communications: An International Journal, 17(2), 98–112.

Asquith, P., Mikhail, M., Au, A., 2005. Information content of equity analyst reports. Journal of Financial Economics, 75(2), 245–282.

Athanasakou, V., Hussainey, K., 2014. The perceived credibility of forward-looking performance disclosures. Accounting and Business Research. DOI: 10.1080/00014788.2013.867403.

Barber, B., Lehavy, R., NcNichols, M., Trueman, B., 2001. Can investors profit from the prophets? Security analyst recommendations and stock returns. Journal of Finance, 65 (2), 531–563.

Barker, R., 1998. The market for information-evidence from finance directors, analysts and fund managers. Accounting and Business Research, 29 (1), 3–20.

Barron, O., Byard, D., Kile, C., Riedl, E., 2002. High-technology intangibles and analysts' forecasts. Journal of Accounting Research, 40 (2), 289–312.

Barron, O., Kile, C.,O'Keefe, T., 1999. MD&A quality as measured by the SEC and analysts' earnings forecasts. Contemporary Accounting Research, 16(1), 75–109.

Barry, C., Brown, S. 1985. Differential information and security market equilibrium. Journal of Financial and Quantitative Analysis, 20 (4): 407–422.

Barth, M., Clinch, G., 1998.Revalued financial, tangible, and intangible assets: associations with share prices and non-market based value estimates. Journal of Accounting Research, 36 (3), 199–233.

Basu, S., Hwang, L.S., Jan, C-L, 1998. International variation in accounting measurement rules and analysts' earnings forecast errors. Journal of Business and Finance and Accounting, 25 (9–10), 1207–1247.

Berthelot, S., Cormier, D., Magnan, M., 2003. Environmental disclosure research: review and synthesis. Journal of Accounting Literature, 22, 1–44.

Beyer, A., Cohen, D., Lys, T., Walther, B., 2010. The financial reporting environment: Review of the recent literature. Journal of Accounting & Economics, 50(2/3), 296–343.

Bolliger, G., 2004, The characteristics of individual analysts' forecasts. Journal of Banking and Finance, 28 (9), 2283–2309.

Bozzolan, S., Trombetta, M., Beretta, S., 2009. Forward-looking disclosures, financial verifiability and analysts' forecasts: A study of cross-listed European firms. European Accounting Review, 18(3), 435–473.

Bradshaw, M., 2004. How do analysts use their earnings forecasts in generating stock recommendations. The Accounting Review, 79 (1), 25–50.

Brammer, S., Pavelin, S., 2008. Factors influencing the quality of corporate environmental disclosure. Business Strategy and the Environment, 17(2), 120–136.

Breton, G., Taffler, R., 2001. Accounting information and analyst stock recommendation decisions: a content analysis approach. Accounting and Business Research, 31(2), 91–102.

Brown, L., Call, A., Clement, M., Sharp, N. 2014. Inside the "black box" of sell-side financial analysts. SSRN Working paper, 66p.

Brown, N., Deegan, C., 1998. The public disclosure of environmental performance information – a dual test of media agenda setting theory and legitimacy theory. Accounting and Business Research, 29(1), 21–41.Bushee, B., Leuz, C., 2005. Economic consequences of SEC disclosure regulation: evidence from the OTC Bulletin Board. Journal of Accounting and Economics, 39 (2), 233–264.

Bushman, R., Smith, A., 2001. Financial accounting information and corporate governance. Journal of Accounting and Economics, 32 (1–3), 237–333.

Campbell, D., Slack, R., 2011. Environmental disclosure and environmental risk: Sceptical attitudes of UK sell-side bank analysts. The British Accounting Review, 43(1), 54–64.

Chen, C., 2009. Who really matters? Revenue implications of stakeholder satisfaction in a health insurance company. The Accounting Review 84 (6), 1781–1804.

Chen, X., Cheng, Q., Lo, K., 2010. On the relationship between analyst reports and corporate disclosures: Exploring the roles of information discovery and interpretation. Journal of Accounting and Economics, 49(3), 206–226.

Cheng, Y., Liu, M., Qian, J., 2006. Buy-side analysts, sell-side analysts, and investment decisions of money managers. Journal of Financial and Quantitative Analysis, 41(1), 51–83.

Chung, K, Jo, H., 1996. The impact of security analysts' monitoring and marketing functions on the market value of firms. Journal of Financial and Quantitative Analysis, 31 (4): 493–512.

Chung, R., Firth, M., Kim, J.-B., 2005. FCF agency costs, earnings management, and investor monitoring. Corporate Ownership and Control, 2 (4), 51–61.

Clement, M., Rees, L., Swanson, E., 2003. The influence of culture and corporate governance on the characteristics that distinguish superior analysts. Journal of Accounting, Auditing and Finance, 18 (4), 593–618.

Cohen, J. R., Holder-Webb, L., Nath, L., Wood, D., 2012. Corporate reporting of nonfinancial leading indicators of economic performance and sustainability. Accounting Horizons, 26(1), 65–90.

Cohen, J., Holder-Webb, L., Nath, L., Wood, D., 2011. Retail investors' perceptions of the decision-usefulness of economic performance, governance, and corporate social responsibility disclosures. Behavioral Research in Accounting, 23(1), 109–129.

Cohen, J., Krishnamoorthy, G., Wright, A., 2008. Waste is our business inc.: The importance of non-financial information in the audit planning process? Journal of Accounting Education, 26(3), 166–178.

Cole, V., Branson, J., Breesch, D., 2012. In search of the invisible user of financial statement and his information needs. The (non)sense of different standards for listed and non-listed companies. International Journal of Accounting, Auditing and Performance evaluation, 8(1), 1–23.

Coram, P. J., Mock, T. J., Monroe, G. S., 2011. Financial analysts' evaluation of enhanced disclosure of non-financial performance indicators. The British Accounting Review, 43(2), 87–101.

Cormier, D., Aerts, W., Ledoux, M.-J., Magnan, M., 2009. Attributes of social and human capital disclosure and information asymmetry between managers and investors. Canadian Journal of Administrative Sciences, 26(1): 71–88.

Cormier, D., Magnan, M., 2013. The economic relevance of environmental disclosure and its impact on corporate legitimacy: An empirical investigation. Business Strategy and the Environment, DOI: 10.1002/bse.1829

Covrig, V., Low, B., 2005. The relevance of analysts' earnings forecasts in Japan. Journal of Business Finance and Accounting, 32 (7–8), 1437–1463.

Dempsey, S., Gatti, J., Grinnell, D., Cats-Baril, W., 1997. The use of strategic performance variables as leading indicators in financial analysts' forecasts. Journal of Financial Statement Analysis, 2 (4), 61–79.

Dhaliwal, D. S., Li, O. Z., Tsang, A., Yang, Y. G., 2011. Voluntary nonfinancial disclosure and the cost of equity capital: The initiation of corporate social responsibility reporting. The Accounting Review, 86(1), 59–100.

Dhaliwal, D.S., Radhakrishnan, S., Tsang, A., Yong George, Y., 2012. Nonfinancial disclosure and analyst forecast accuracy: International evidence on corporate social responsibility disclosure. The Accounting Review, 87(3), 723–759.

Dickins, D., Higgs, J., 2005. Interpretation and use of auditor fee disclosures. Financial Analysts Journal, 61 (3), 96–102.

Dogl, C., Holtbrugge, D., 2014. Corporate environmental responsibility, employer reputation and employee commitment: an empirical study in developed and emerging economies. International Journal of Human Resource Management, 25(12), 1739–1762.

Dontoh, A., Radhakrishnan, S., Ronen, J., 2004. The declining value-relevance of accounting information and non-information-based trading: An empirical analysis. Contemporary Accounting Research, 21(4), 795–812.

Doukas, J., Kim, C., Pantzalis, C., 2000. Security analysis, agency costs, and company characteristics. Financial Analysts Journal, 56 (6), 54–63.

Durnev, A., Kim, E., 2005. To steal or not to steal: firm attributes, legal environment and valuation. Journal of Finance, 60 (3), 1461–1493.

Eccles, R.G., Serafeim, G., Krzus, M.P., 2011. Market interest in nonfinancial information. Journal of Applied Corporate Finance, 23(4), 113–127.

Elgers, P., Lo, M., Pfeiffer, R., 2001. Delayed security price adjustments to financial analysts' forecast of annual earnings. The Accounting Review, 76 (4), 613–632.

Ellis, J., Fee, E., Thomas, S., 2012. Proprietary costs and the disclosure of information about customers. Journal of Accounting Research, 50(3), 685–727.

European Commission (EC), 2011a, Communication from the commission to the European Parliament, the Council, the Economic and Social Committee and the Committee of the Regions – Single Market Act – Twelve levers to boost growth and strengthen confidence "Working together to create growth". COM (2011) 206, Brussels.

European Commission (EC), 2011b. Summary report on the responses received to the public consultation on disclosure of non-financial information by companies. Brussels.

European Commission (EC), 2014. Non-Financial Reporting. available at: http://ec.europa.eu/internal_market/accounting/non-financial_reporting/ (consulted 22 April 2014)

Fernández, B., Callen, Y., Gadea, J., 2011. Stock price reaction to non-financial news in European technology companies. European Accounting Review, 20(1), 81–111.

Financial Accounting Standards Board (FASB), 2001. Improving business reporting: Insights into enhancing voluntary disclosures. Steering Committee Report, Business Reporting Research Project, 90p.

Flöstrand, P., 2006. The sell side-observations on intellectual capital indicators. Journal of Intellectual Capital, 7(4), 457–473.

Flöstrand, P., Ström, N., 2006. The valuation relevance of non-financial information. Management Research News, 29(9), 580–597.

Fogarty, T., Rogers, R., 2005. Financial analysts' reports: an extended institutional theory evaluation. Accounting, Organizations and Society, 30(4), 331–356.

Francis J., Schipper, K., 1999. Have financial statements lost their relevance. Journal of Accounting Research, 37 (2), 319–352.

Francis, J., Schipper, K., Vincent, L., 2002. Earnings announcements and competing information. Journal of Accounting and Economics, 33 (3), 313–342.

García-Meca, E., 2005. Bridging the gap between disclosure and use of intellectual capital information. Journal of Intellectual Capital, 6(3), 427–440.

García-Meca, E., Martinez, I., 2007. The use of intellectual capital information in investment decisions – an empirical study using analyst reports. International Journal of Accounting, 42(1), 57–81.

García-Meca, E., Parra, I., Larran, M., Martinez, I., 2005. The explanatory factors of intellectual capital disclosure to financial analysts. European Accounting Review, 14 (1), 63–94.

Gietzman, M., Ireland, J., 2005. Cost of capital, strategic disclosures and accounting choice. Journal of Business Finance and Accounting, 32 (3–4), 599–634.

Gomes, A., Gorton, G., Madureira, L., 2007. SEC regulation fair disclosure, information, and the cost of capital. Journal of Corporate Finance, 13 (2–3), 300–334.

Groysberg, B., Healy, P., Chapman, C., 2008. Buy-side vs sell-side analysts' earnings forecasts. Financial Analysts Journal, 64 (4), 25–39.

Hail, L., 2013. Financial reporting and firm valuation: relevance lost or relevance regained? Accounting and Business Research, 43(4), 329–358.

Hirst, E., Koonce, L., Simko, P., 1995. Investor reactions to financial analysts' research reports. Journal of Accounting Research, 33 (2), 335–351.

Hoff, B., Wood, D., (2008). The use of non-financial information: What do investors want?. Project Findings Boston College Carroll School of Management, 16p.

Holder-Webb, L., Cohen, J., Nath, L., Wood, D., 2009. The supply of corporate social responsibility disclosures among U.S. Firms. Journal of Business Ethics, 84(4), 497–527.

Holland, J., 2003. Intellectual capital and the capital market – organisation and competence. Accounting, Auditing and Accountability Journal, 16 (1), 39–47.

Holland, J., Johanson, U., 2003. Value-relevant information on corporate intangibles – creation, use and barriers in capital markets – "between a rock and hard place". Journal of Intellectual Capital, 4(4), 465–486.

Hong, H., Kubik, J., Salomon, A., 2000. Security analysts' career concerns and the hearding of earnings forecasts. Rand Journal of Economics, 31(1), 121–144.

Hope, O.-K., Thomas, W. B., Winterbotham, G., 2006. The Impact of nondisclosure of geographic segment earnings on earnings predictability. Journal of Accounting, Auditing & Finance, 21(3), 323–346.

Hope, O-K, 2003. Analyst following and the influence of disclosure components, IPOs and ownership concentration. Aisa-Pacific Journal of Accounting and Economics, 10 (2), 117–141.

Hsu, W., Chang, Y., 2011. Intellectual capital and analyst forecast: evidence from the high-tech industry in Taiwan. Applied Financial Economics, 21(15), 1135–1143.

Hunt, H. G., Grinnel, D. J., 2003. Financial analysts' view of the value of environmental information. Advances in Environmental Accounting & Management, 2, 101–120.

Hutton, A., Miller, G.S., Skinner, D.J., 2003. The role of supplementary statements with earnings forecasts. Journal of Accounting Research, 41(5), 867–890.

International Federation of Accountants (IFAC), 2013. Enhancing organizational reporting. IFAC Policy Position 8, New York, 8p.

Ivković, Z., Jegadeesh, N., 2004. The timing and value of forecast and recommendation revisions. Journal of Financial Economics, 73(3), 433–463.

Jiraporn, P., Gleason, K., 2007. Delaware incorporation and earnings management: an empirical analysis. Journal of Applied Finance, 17(1), 40–51.

Johanson, U., 2003. Why are capital markets actors ambivalent to information about certain indicators on intellectual capital. Accounting, Auditing and Accountability Journal, 16 (1), 31–38.

Johansson, J., 2007. Sell-side analysts' creation value – key roles and relational capital. Journal of Human Resource Costing and Accounting, 11(1), 30–52.

Juntilla, J., Kallunki, J.-P., Kärja, A., Martikainen, M., 2005. Stock market response to analysts' perceptions and earnings in a technology-intensive environment. International Review of Financial Analysis, 14(1), 77–92.

Kallapur, S., Kwan, S., 2004. The value relevance and reliability of brand assets recognized by U.K. firms. The Accounting Review, 79(1), 151–172.

Khlifi, F., Bouri, A., 2010. Corporate disclosure and firm characteristics: A puzzling relationship. Journal of Accounting, Business & Management, 17(1), 62–89.

Khurana, I., Raman, K., 2004. Litigation risk and the financial reporting credibility of Big 4 versus non-Big 4 audits: evidence from Anglo-American countries. The Accounting Review, 79(2), 473–495.

LaFond, R., Watts, R., 2008. The information role of conservatism. The Accounting Review, 83(2), 447–478.

Lang, M., Lundholm, R., 1996. Corporate disclosure policy and analyst behavior. The Accounting Review, 71(4), 467–492.

Leuz, C., Verrecchia, R., 2000. The economic consequences of increased disclosure. Journal of Accounting Research, 38(3), 91–124.

Lev, B., Zarowin, P., 1999. The boundaries of financial reporting and how to extend them. Journal of Accounting Research, 37 (2), 353–385.

Lev, B., 2004. Sharpening the intangibles edge. Harvard Business Review, June, 109–116.

Liang, C.-J., Yao, M-L., 2005. The value relevance of financial and nonfinancial information – evidence from Taiwan's information electronics industry. Review of Quantitative Finance and Accounting, 24 (2), 135–157.

Livnat, J., Zhang, Y., 2012. Information interpretation or information discovery: which role of analysts do investors value more? Review of Accounting Studies, 17(3), 612–641.

Luo, X., Homburg, C., Wieseke, J., 2010. Customer satisfaction, analyst stock recommendations, and firm value. Journal of Marketing Research, 47(6), 1041–1058.

Luo, X., Wang, H., Raithel, S., Zheng, Q., 2014. corporate social performance, analyst stock recommendations, and firm future returns. Strategic Management Journal. DOI: 10.1002/smj.2219.

Maines, L., Bartov, E., Fairfield, P., Hirst, D., Iannaconi, P., Mallet, R., Schrand, C., Skinner, D., Vincent, L., 2002. Recommendations on disclosure of nonfinancial performance measures. Accounting Horizons, 16(4), 353–362.

Maines, L., McDaniel, L., 2000. Effects of comprehensive-income characteristics on nonprofessional investors' judgments: The role of financial-statement presentation format. The Accounting Review, 75 (2), 179–207.

Marston, C., Polei, A., 2004. Corporate reporting on the internet by German companies. International Journal of Accounting Information Systems, 5 (3), 285–311.

Matsumoto, D., Pronk, M., Roelofsen, E., 2011. What makes conference calls useful? The information content of managers' presentations and analysts' discussion sessions.The Accounting Review, 86(4), 1383–1414.

Mohd Ghazali, N. A., Weetman, P., 2006. Perpetuating traditional influences: Voluntary disclosure in Malaysia following the economic crisis. Journal of International Accounting, Auditing and Taxation, 15(2), 226–248.

Moser, D., Martin, P., 2012. A broader perspective on corporate social responsibility research in accounting. The Accounting Review, 87(3), 797–806.

Ngobo, P.-V., Casta, J.-F., Ramond, O., 2012. Is customer satisfaction a relevant metric for financial analysts? Journal of the Academy of Marketing Science, 40(3), 480–508.

Nichols, C., Wieland, M., 2009. Do firms' non-financial disclosures enhance the value of analyst services? Working paper Cornell University, 44p.

Nielsen, C., 2008. A content analysis of analyst research: health care through the eyes of analysts. Journal of Health Care Finance, 34(3), 66–90.

NIVRA, 2010. Niet-financiële informatie steeds belangrijker voor beleggers, beschikbaar op www.accountancynieuws.nl/actueel/accountancymarkt/niet-financiele-informatie-van-beursfondsen-steeds.92618.lynkx, consulted september 2012.

Orens, R., Aerts, W., Cormier, D., 2010. Web-based non-financial disclosure and cost of finance. Journal of Business Finance & Accounting, 37(9/10), 1057–1093.

Orens, R., Lybaert, N., 2007. Does the financial analysts' usage of non-financial information influence the analysts' forecast accuracy? Some evidence from the Belgian sell-side financial analyst. International Journal of Accounting, 42(3), 237–271.

Orens, R., Lybaert, N., 2010a. Niet-financiële informatierapportering en de kwaliteit van de winstprognoses opgesteld door financiële analisten. Accountancy en Bedrijfskunde, 30(6), 3–15.

Orens, R., Lybaert, N., 2010b. Determinants of sell-side financial analysts' use of non-financial information. Accounting and Business Research, 40(1), 39–53.

Pinho, C., Madaleno, M., Santos, H., 2013. The usefulness of financial analysts' reports: A content analysis. International Journal of Management, 31(2), 631–648.

Prencipe, A., 2004. Proprietary costs and determinants of voluntary segment disclosure: evidence from Italian listed companies. European Accounting Review, 13 (2), 319–340.

Previts, G. J., Bricker, R. J., Robinson, T. R., Young, S. J., 1994. A content analysis of sell-side financial analyst company reports. Accounting Horizons, 8(2), 55–70.

Ramnath, S., Rock, S., Shane, P., 2008. The financial analyst forecasting literature: A taxonomy with suggestions for further research. International Journal of Forecasting, 24(1), 34–75.

Rogers, R. K., Grant, J., 1997. Content analysis of information cited in reports of sell-side financial analysts. Journal of Financial Statement Analysis, 3(1), 14–30.

Said, A., Hassab Elbany, H., Wier, B., 2003. An empirical investigation of the performance consequences of nonfinancial measures. Journal of Management Accounting Research, 15, 193–223.

Sakakibara, S., Hansson, B., Yosano, T., Kozumi, H., 2010. Analysts' perceptions of intellectual capital information. Australian Accounting Review, 20(3), 274–285.

Schipper, K., 1991. Commentary on analysts' forecasts. Accounting Horizons, 5(4), 105–121.

Schuler, D. A., Cording, M., 2006. A corporate social performance – Corporate financial performance behavioral model for consumers. The Academy of Management Review, 31(3), 540–558.

Simnett, R., Vanstraelen, A., Chua, W., 2009. Assurance on sustainability reports: An international comparison. The Accounting Review, 84(3), 937–967.

Simpson, A. N., 2010. Analysts' use of nonfinancial information disclosures. Contemporary Accounting Research, 27(1), 249–288.

Skinner, D., 2008. Accounting for intangibles – a critical review of policy recommendations. Accounting and Business Research, 38 (3), 191–204.

Soliman, M., 2013. Firm characteristics and the extent of voluntary disclosure: the case of Egypt. Research Journal of Finance and Accounting, 4(17), 71–81.

Stark, A., 2008. Intangibles and research – an overview with a specific focus on the UK. Accounting and Business Research, 38(3), 275–285.

Turban, D. B., Greening, D. W., 1997. Corporate social performance and organizational attractiveness to prospective employees. The Academy of Management Journal, 40(3), 658–672.

Upton, W., 2001. Special report on business and financial reporting, challenges from the new economy. Financial Accounting Standards Board, Norwalk, Connecticut, 118p.

Vanstraelen, A., Zarzeski, M., Robb, S., 2003. Corporate nonfinancial disclosure practices and financial analyst behaviour across three European countries. Journal of International Financial Management and Accounting, 14(3), 249–278.

Yu, M., 2010. Analyst forecast properties, analyst following and governance disclosures: A global perspective. Journal Of International Accounting, Auditing & Taxation, 19(1), 1–15.

Zéghal, D., Maaloul, A., 2011. The accounting treatment of intangibles – A critical review of the literature. Accounting Forum, 35(4), 262–274.

THE CONTRIBUTION OF DYNAMIC CAPABILITIES TO NEW VENTURE SURVIVAL IN NASCENT MARKETS: THE BOUNDARY ROLE OF STABILITY

Robin De Cock* and Bart Clarysse

Abstract

The dynamic capability literature has argued that dynamic capabilities are of most importance to companies that face dynamic environments. New ventures in nascent markets are in such a situation. They need to develop dynamic capabilities to survive. However, the literature remains silent when it comes to the boundary conditions under which these dynamic capabilities have most impact on survival. We extend the literature on dynamic capabilities by arguing that firm stability triggered by the formalization of the roles in the management team and the installation of an external board will help the firm to organize and structure the key organizational resources around a particular opportunity and subsequently facilitates and increases the impact of dynamic capabilities. We therefore contribute to the literature on dynamic capabilities by showing its duality with firm stability. However, while role formalization is an important boundary condition for the success of dynamic capabilities, the installation of a board with external directors decreases the impact of dynamic capabilities. We explain this by the decreased resource cognition among the decision makers at board level which undermines the positive impact of venture stability.

Keywords: dynamic capabilities; firm survival; firm stability; nascent markets

JEL codes: L1, L26, M13

* Corresponding author: Robin De Cock, UGent, Department of Management, Innovation and Entrepreneurship, Tweekerkenstraat 2, 9000 Gent. Email: Robin.DeCock@UGent.be.

I. INTRODUCTION

The extant literature on dynamic capabilities has focused on how large companies gain a competitive advantage if they are able to develop these capabilities in preferably but not exclusively dynamic environments (Helfat and Peteraf, 2003; Teece, 2007; Helfat and Winter, 2011). Dynamic capabilities are generally considered as the capacity of a firm to change its resource base (Helfat and Peteraf, 2009). Teece (2007) further defined dynamic capabilities as the capacity of a firm (1) to sense and shape opportunities and threats, and (2) to seize opportunities and (3) reconfigure the existing firm's resources. As such, dynamic capabilities are necessary to change the existing key organizational resources towards new opportunities. Zahra, Sapienza, & Davidsson (2006) extended the dynamic capability perspective to entrepreneurial companies, arguing that also these companies benefit from dynamic capabilities as they allow new ventures to be able to continuously create, define, discover and exploit entrepreneurial opportunities. Dynamic capabilities differ from entrepreneurial capabilities as they encompass the firm's capacity to change its business model and resource base towards new, emerging opportunities whereas entrepreneurial capabilities refer to the identification of opportunities and the development of a resource base to pursue these opportunities (Arthurs and Busenitz, 2006). Overall, the dynamic capability perspective suggests that dynamic capabilities are needed to build up a competitive advantage both in large and small firms. However, the extant literature falls short in explaining the boundary conditions at firm level under which these dynamic capabilities lead to better performance. This is the theoretical gap we address in this paper.

Researchers in the domain originally assumed that dynamic environments triggered the use of a firm's dynamic capabilities (Teece, Pisano, Shuen, 1997; Teece, 2007). More recently, Helfat and Winter (2011) and Drnevich and Kriauciunas (2012) showed that although dynamic capabilities might be more useful in dynamic environments, they also lead to competitive advantage in less dynamic ones. Beyond the environment, the literature remains silent when it comes to boundary conditions at company level. Zahra et al. (2006) hypothesize that dynamic capabilities will accrue over time and form a complex set of inter-relations with operational or substantive capabilities but do not touch upon the internal organizational form which might be needed to optimize the impact of dynamic capabilities. Farjoun (2010) argues that change is most effective when it is embedded within stability. This implies that dynamic capabilities might be most effective in an organization which also has sufficient stability to embody change. Even in new ventures, stability is needed to avoid the chaos which tends to be associated with major changes (D'Aveni, Dagnino and Smith, 2010). In other words, changing your resource base and jumping from one opportunity to another, doesn't necessarily lead to new successes. New ventures need to create stability by installing mechanisms that organize and manage the key organizational resources around the new opportunities that are either identified or shaped by the new venture (Sirmon et al, 2007).

These mechanisms should support the new venture to (re)structure the firm's resource portfolio, bundle the resources to build strong substantive capabilities, and leverage those capabilities with the purpose of creating and maintaining value for customers and owners (Sirmon et al., 2007). As such, the literature on resource management (Sirmon et al, 2007; 2011) helps us to explain how change can be facilitated by mechanisms that reinforce stability.

Based on the extant entrepreneurship literature, we have identified two important mechanisms which could support these resource management processes (Sirmon et al., 2007) and lead to stability in new ventures: (1) the formalization of the roles in the management team and (2) the composition of an external board. These factors are even more important for the stability of new ventures active in nascent markets. Nascent markets are environments that are characterized by turbulence and uncertainty and therefore necessitate change and flexibility (Sine and David, 2003; Santos and Eisenhardt, 2009). First, Sine, Mitsuhashi and Kirsch (2006) show that new ventures in such markets need formal structures in order to overcome liabilities of newness (Stinchcombe, 1965). Whereas formal structures tend to inhibit change in large, established firms, new ventures typically start-up in dynamic environments and need a structured founding or (early) top management team to deal with the role ambiguity and uncertainty which accompanies change in the environment (Sine et al., 2006). Second, new ventures reach stability through the involvement of external stakeholders which act as financers, catalysts and monitors (Garg, 2012). External stakeholders tend to participate in the decision making process through the board of directors which forms the key governance mechanisms of new ventures (Dowell and Schackell, 2011). Boards track the significant behaviors of the founders, the outcomes of their actions, and the performance of the venture in order to ensure that corrective action is taken as needed (Garg, 2012). Monitoring in new ventures increases their stability in markets that call for frequent changing of direction because changes are extensively discussed and benchmarked before implementation. When the firm decided to restructure their business activities, boards use their experience to advise the management team and provide them access to the necessary resources to support these changes. In this paper, we empirically test whether dynamic capabilities will decrease the probability of failure of new ventures in nascent markets. We theoretically contribute to the dynamic capability literature by extending this perspective into the extant literature on organizational theory which proposes organizational stability as an important determinant of firm success and which has recently shed a new light on the relation between stability and change, presenting it as a duality of reinforcing concepts (Farjoun, 2010). We therefore hypothesize that organizational stability is a boundary condition for dynamic capabilities to enhance the survival potential of a new venture.

To examine our central questions and to test our hypotheses, we use a panel of 124 new ventures founded in the period 2006–2008, which we followed over time in the period 2009–2012 using several interview rounds to collect data on the development of their dynamic

capabilities. We used a hazard model to investigate the impact of dynamic capabilities on firm survival. Hazard models have been used extensively and in a wide variety of contexts in the innovation and strategy literatures (Sinha and Noble, 2008). This type of analysis allows for the modeling of failure at each time point, and considers both the occurrence and timing of a failure (Cui, Calantone, & Griffith, 2010). The new ventures were selected based upon the fact that they did apply for an innovation grant with the Flemish Government to finance the development of a business plan with the specific objective to raise venture capital. Only new ventures which enter into nascent markets of which the industry structure is not clear yet are eligible for this type of financing. We choose this empirical context because new ventures are less complex than more established firms, so they provide a comparatively clean setting for an empirical exploration of the effect of resources and capabilities (Gruber, Heinemann, Brettel, & Hungeling, 2010).

This study theoretically extends the literature on dynamic capabilities towards organization theory where scholars tend to focus on stability as an important element of organizational performance. First, we show that formalization improves the impact of dynamic capabilities on the performance of new ventures. More specifically, we show that dynamic capabilities benefit from clear internal role specialization and formalization of the founding team. Second, we show that boards, which are considered a second source of stability, do not have the same impact on the relation between dynamic capabilities and new venture performance. On the contrary, boards limit the impact of dynamic capabilities. Boards typically monitor the new venture performance based upon an agreed business plan which is difficult to change. External directors in boards might be too distant from the new venture's operations to be assistive in implementing changes.

We structure the paper as follows. First, we draw on theoretical insights from both organizational design and boards as well as dynamic capabilities to develop our hypotheses. Next, we present details on our methodological and sampling approach. Finally, we close with a discussion of the results and implications for theory and practice.

II. THEORY AND HYPOTHESES DEVELOPMENT

New ventures in emergent economic sectors or nascent markets start up in environments that are characterized by turbulence and uncertainty (Sine and David, 2003; Santos and Eisenhardt, 2009). Nascent markets lack a dominant logic to guide actions (Kaplan and Tripsas, 2008) and therefore form important challenges to new ventures which operate in these markets. Due to a lack of legitimated industry logics, the new ventures have difficulties to identify which resources are strategic (Bingham, Eisenhardt & Davis, 2009) and to develop associated business models (Ozcan and Eisenhardt, 2009). Survival rather than efficiency is the main objective of ventures in these markets (Santos and Eisenhardt, 2009). New ventures that want to survive in these nascent markets will need to develop capabilities

which allow them to navigate through the ambiguity which usually is associated with these markets.

The capability literature has made a distinction between substantive and dynamic capabilities (Eisenhardt and Martin, 2000; Helfat and Peteraf, 2003; Zahra et al., 2006; Helfat and Winter, 2011). Whereas substantive capabilities represent the firm's capacity to develop routines which make its living and hence directly contribute to the efficiency of the company, dynamic capabilities refer to its capacity to change its resource configuration and business model. Teece (2007) further describes a firm's dynamic capabilities as its capacity to 'sense' opportunities, 'seize' these opportunities in terms of developing an appropriate business model and eventually 'implement' change through applying this business model.

Early proposals in this field clearly assumed a direct relationship between firms' dynamic capabilities and their performance (Teece et al., 1997). These authors stated that this framework is intended to explain firm-level success and failure, competitive advantage, and private wealth creation (Teece et al., 1997; Makadok, 2001; Zollo and Winter, 2002). More recently, Teece (2007) stated that "the ambition of the dynamic capabilities framework is nothing less than to explain the sources of enterprise-level competitive advantage over time" and that "dynamic capabilities lies at the core of enterprise success (and failure)." However, other researchers took a more cautious approach towards the relation between performance and dynamic capabilities. In their view, long-term competitive advantage does not only rely on dynamic capabilities themselves but on the key resource configurations or substantive capabilities created by the dynamic capabilities (Eisenhardt and Martin, 2000; Winter, 2003; Helfat and Winter, 2011). They argue that a firm's resource base include key organizational resources that are efficiently combined, bundled and structured by the firm's substantive capabilities and subsequently have a direct effect on the firm performance However, these theoretical arguments are mostly developed in the context of established firms and for the purpose of creating a sustainable competitive advantage (Barreto, 2010).

Since substantive capabilities are efficiency oriented, we might expect that dynamic capabilities will be the key capabilities that are needed to survive the difficult early stages of new ventures in nascent markets. The development of dynamic capabilities will allow the new venture to be alert for new emerging logics in the market and will allow it to adjust its business model accordingly, if necessary. On the contrary, the lack of dynamic capabilities might constrain the viability of a new venture in a nascent market. Bingham (2009) shows that experimentation is needed in seizing the opportunities in order to be successful in new, unfamiliar markets. This means that new ventures need to be able to experiment with different business models and resource configurations that are in line with these business models in order to be successful in these markets. Hence, a lack of dynamic capabilities, which allow new ventures to detect new logics in the market and eventually adjust their business model and resource configuration, will lead to rigidities and eventually lead to new

venture failure. The key resources will be outdated and not useful anymore in the light of emerging new opportunities and changing environments that characterizes nascent markets. Autio, George and Alexy (2011) argue that in environments such as nascent markets a lack of substantive capabilities might even be an advantage. This implies that those ventures which develop substantive capabilities will even have higher failure rates if they lack dynamic capabilities which allow them to change these substantive capabilities than if they have no capabilities at all. In line with the dynamic capability literature, we therefore hypothesize:

Hypothesis 1. The Development of Dynamic Capabilities will decrease the Probability of Failure of New Ventures in Nascent Markets

While dynamic capabilities can be viewed as important mechanisms to guide new ventures through the difficult early stages, Farjoun (2010) does alert that in order to survive, organizations must reconcile stability with change. Organizing for firm survival and growth in nascent markets means that new ventures should be able to sense and seize new opportunities and reconfigure existing resources and capabilities (Teece, 2007). However, the level of rivalry and innovativeness in these nascent markets could escalate, making dynamic capabilities the instrument of ever greater chaos (D'Aveni et al., 2010). Therefore, organizational behaviorists suggests that firms need "stable building blocks" in order to facilitate change and benefit from these changes (Farjoun, 2010; Schreyögg and Sydow, 2010; Smith and Lewis, 2011). These studies advocate structure and stability as necessary elements to undertake change. New ventures can create stability through the resource management processes of structuring and bundling their key resources around new opportunities identified or shaped through their dynamic capabilities.

Along the same lines, Sine et al. (2006) already emphasized that new ventures in nascent markets need formal structure to prosper in these markets and overcome what Stinchcombe (1965) has referred to as the liability of newness which new ventures face. This implies that in new ventures, especially in nascent markets, some form of structure is needed in which dynamic capabilities should be embedded. Zahra et al. (2006: 918) argue that Teece's organizational level process of sensing, seizing and shaping opportunities corresponds in new ventures to the entrepreneur, the entrepreneurial team or the firm's senior management 'perception' of opportunities, their 'willingness' to undertake change and their 'ability' to implement changes. In other words, a central role is allocated to the founding team and the key decision makers in the dynamic capability process. However, the dynamic capability literature remains largely silent when it comes to describe how these founding teams might impact the overall relation between dynamic capabilities and performance (Sirmon et al., 2011).

Sine et al. (2006) formalize structure by identifying role formalization in founding teams. Following Dalton, Todor, Spendolini, Fielding and Porter (1980), Sine et al. (2006:122)

define formalization of organizational tasks as the 'identification and designation of particular functional roles and their assignment to specific individuals'. Role formalization avoids confusion about who is supposed to do particular routine tasks. Having formalized roles in the founding team of a new venture implies that there is a clear attempt to decrease the ambiguity of the environment as each team member will know exactly what to do and coordination costs decrease. Coordination costs refer to the costs associated with the efforts needed to resolve disputes, disagreements, or conflicts about the nature and the scope of the change needed (Zahra and Nielsen, 2002). Zahra et al. (2006) already refer to the need for integration if dynamic capabilities need to be developed. Dynamic capabilities assume that opportunities are identified or shaped and decisions are made about how to address new, emerging opportunities. If the roles in the founding team are clearly allocated, it is likely that these decisions will be taken easier since every member has a specific domain (e.g. technology, marketing, operations, …) for which he/she is responsible and trustworthy. A formalized team will be capable of structuring the key resources, bundling these resources into substantive capabilities in an efficient way (Sirmon et al., 2007) and aligning them with the new opportunity space that is shaped or identified through the firm's dynamic capabilities. As new ventures in nascent markets are confronted with changing environments, formalized teams will be faster than other teams in facilitating the process of change initiated by the firm's dynamic capabilities through a more efficient key resource management.

A lack of role formalization might on the contrary lead to total chaos in the case of change. In the latter case, the different founding members will have an opinion about all the functional domains and about what needs to be done in each of these domains in order to adjust to the new insights or opportunities which emerge when markets develop. A lack of clear role alienation will force new ventures to rely upon decision making by consensus and will slow down the resource management processes necessary to develop substantive capabilities needed to exploit the new opportunities and reinforce the effect of dynamic capabilities (Sine et al., 2006). Unclear roles increases the time and costs to arrive at any particular decision which destroys the effect of the firm's dynamic capabilities as other players in the nascent market could already started to exploit the new opportunities and begin to conquer the market resulting in high market shares. In other words, developing dynamic capabilities and getting most out of them will become extremely difficult in these ventures. At any of the three stages in the process of these capabilities, a lack of consensus can collapse the impact of capabilities. We therefore hypothesize:

Hypothesis 2. The negative impact of dynamic capabilities on the probability of failure will increase with greater role formalization in the founding team

Sine et al. (2006) refer to the role formalization among founding team members as an indicator of organizational structure and flexibility. However, not only internal structures do

bring stability to new ventures. Garg (2012) argues that the boards of directors in these ventures are of critical importance because they have, as a key governance mechanism in new ventures, a monitoring function in addition to their more frequently recognized advisory role (Wasserman and Boeker, 2010). Monitoring, which can be defined as the director's activities which involve the tracking of founder behavior to make sure that corrective action is taken if needed, is critical to ensure the stability of ventures in markets that call for frequent changing of directions such as nascent markets. The monitoring function of boards in new ventures is distinct from public firms, where boards almost exclusively have been studied, as the separation between ownership and control disappears in new ventures. The key management typically consists of the founding team, which tends to have similar financial interests with other firm owners such as outside investors that are represented in the board (Wasserman, 2006). Because of the financial stake of these investors in the new ventures, the latter tend to be more involved in monitoring than typical directors in public firms. Hence, one can see the board of directors of a new venture as the enlarged management team which monitors the actions of the founder-managers.

As new ventures in nascent markets do face ambiguity (Santos and Eisenhardt, 2009) and shifting industry structures (Ozcan and Eisenhardt, 2009), traditional financial metrics such as profit are usually unavailable. Therefore venture board members will monitor both strategic and operational activities and will do this on a frequent basis. Typically, venture boards are likely to make sure that strategic decisions of the founder-CEOs keep the firm's focus on growth instead of the personal goals of these founder-CEOs such as realizing an original product vision and maintaining a particular organizational culture (Tuggle, Sirmon, Reutzel, & Bierman 2010; Garg, 2012). Therefore, it is likely that boards will both stimulate and challenge new opportunities that emerge and/or new logics that are formed in nascent markets and will advise the founding team on how to structure and bundle their key resources around the new opportunities. Moreover, besides their monitoring and advisory role, boards can also provide new ventures the access to resources to adjust the new venture's business model in line with these new identified opportunities or new formed logics (Dowell and Schackell, 2011). In sum, boards provide a formal structure to the new venture which allows the founding team to benchmark its ideas, forces the team to carefully reflect upon potential changes and gives them access to additional resources.

Despite the fact that boards facilitate structure and financial reporting procedures, they might also invoke rigidity in a company because of resource cognition (Danneels, 2010). Resource cognition refers to the cognitions which managers have about the firm's resources. More specifically, resources cognition refers to the identification of resources and the understanding of their fungibility, which is crucial in understanding the impact of dynamic capabilities. Danneels (2010) shows that the further away executive decision makers are from the work floor the more difficult it will be for them to assess the real resources of the company. Hence, directors in a venture might rely on cognitions which are

detached from the real environment in which the venture operates or which are different from the real resource base that can form a competitive environment. We can imagine that directors, who are only occasionally present at the venture and who serve on different venture boards, do not have the same focus of attention as founding team members. They could have different views on the key resource which will have implication for the structuring and bundling processes necessary to facilitate the positive effects of dynamic capabilities. Still, due to the power of boards in ventures, they will codetermine strategic decisions and monitor the actions taken by the venture executives. Their deep involvement in the strategic decision making might be counterproductive, especially in nascent markets where usually no financial performance indicators are possible to use as benchmarks. That's why we believe that the monitoring and advising role of boards will be counterproductive as they do not understand the nascent market as good as the founders do and subsequently have difficulties to identify the key organizational resources. This could inhibit or slow down change stimulated through the firm's dynamic capabilities when this is crucial for survival.

Therefore we hypothesize:

Hypothesis 3. The negative impact of dynamic capabilities on the probability of failure will decrease with the availability of an external board in the new venture

III. METHODS

Our objective in this research is to consider the boundary conditions on the central premise that dynamic capabilities influence new venture survival. We use a hazard modeling framework to investigate the impact of dynamic capabilities on firm survival. Hazard models have been used extensively and in a wide variety of contexts in the innovation and strategy literatures (Sinha and Noble, 2008).

A. SAMPLE AND DATA COLLECTION

This study gathered survey and secondary data on new technology-based firms founded in Flanders (northern part of Belgium). We started with a list of 211 new ventures, which was provided by the Flemish agency for innovation by science and technology (IWT). After checking the founding years from the BELFIRST database, we chose to eliminate companies older than 3 years at the time of the first interview round (2009). This approach is consistent with the operationalization of new ventures (Zahra, Hitt, & Ireland, 2000). Even though different age ranges have been used in the literature, there is a growing consensus that firms 6 years and younger are new ventures (Zahra et al., 2000). In all, we contacted 185 companies. Of these, 6 new ventures had ceased to exist by the time we sought to contact them, because

they had been acquired or dissolved. Eventually, we collected general data about the founders, management teams, business models and founding conditions of 148 new ventures (response rate 80%) in the first interview round. This baseline information was collected by using a structured questionnaire during face-to-face interviews. These face-to-face interviews were extremely helpful to explain the upcoming longitudinal, quantitative research design. Interview duration varied from 30 minutes to two hours. The interviews also allowed us to build trust and rapport with the founders to increase cooperation and response rate during the quantitative, longitudinal data collection process. Interviews with these companies revealed their business model and future plans which confirmed their innovative reputation. All these companies were granted by the IWT because they had developed technological innovations which could have a significant economic impact. Based on the pre-selection by the IWT, the face-to-face interviews and an extensive web search, we concluded that these companies were active in nascent markets. ICT companies in our sample focus for example on mobile internet or cloud software. Engineering companies are developing solutions for electronic vehicles or invented new ways to save and generate energy. Biotech companies in our sample develop new generation of drugs that has the potential to treat a broad range of severe diseases.

In the end, we followed 230 founders of 124 new ventures (response rate of 67%) through the first interview round (2009) and the two follow-up rounds (2010 and 2011). In these follow-up rounds, we collected information about the entrepreneurial team and the capabilities the companies had developed after start-up. Here, we used a web-based survey supported by telephone follow-ups to collect capability data. We added company data on each of the ventures, which we collected from IWT (the innovation granting institute which supported this research), BELFIRST, GRAYDON, VENTUREXPERT and the Belgian Official Journal. By doing so, we managed to collect data on environmental dynamism, firm survival, types of investors, amount of raised capital, revenues, employees, sector, etc. In sum, we use several information sources to collect data on the entrepreneurial teams and their companies.

B. DEPENDENT VARIABLE

Our dependent variable is firm failure. Firms were coded 1 if they failed during the time period studied and 0 otherwise. Failures included completed bankruptcies, completed liquidations, closures based on company request, and merger or acquisition of organizations at risk of bankruptcy (Hannan and Freeman, 1989). We first identified whether a start-up had failed using the Belgian Official Journal. Secondly, we also used financial reports from GRAYDON to identify those companies that are having difficulties to fulfill their financial obligations. The founders of these firms were contacted and coded "1" if the founder confirmed that the company was bankrupt, liquidated or closed. Finally, we also investigated

the rather small amount of companies that were acquired (3%) or were involved in a merger (0.8%). Based on the same financial reports from GRAYDON and reports from VENTUREXPERT, we classified an acquisition as unfavorable using the following criteria: (i) for VC-funded start-ups, if the transaction value (the value of the acquisition deal) was less than the total capital raised; (ii) if a start-up was not VC funded and reported a loss in the year prior to the acquisition; (iii) if the start-up is not VC-funded and we lack profitability data, if none of the founders of the focal start-up joined the acquiring firm (Arora and Nandkumar, 2011).

C. INDEPENDENT VARIABLES

Dynamic capabilities. The capabilities were measured using a scale we developed to capture the extent to which new ventures have the capability to change. The scale was developed based on the scale of Danneels (2008) and the theoretical definition of dynamic capabilities by Teece (2007). Namely, Teece (2007) defines dynamic capabilities as the capacity of a firm (1) to sense and shape opportunities and threats, and (2) to seize opportunities and (3) reconfigure the existing firm's resources. We started by developing an initial pool of scale items based on the scale of Danneels (2008) and the theoretical work of Teece (2007). The initial pool of items was then pre-tested in an interview round with four new technology-based ventures. In each round, two to three interviewees from each venture were asked to complete the questionnaire. While completing the questionnaire, entrepreneurs verbalized any thoughts that came to their minds. The items were revised following each interview round. At the end of round four, feedback from the respondents indicated that the scale items were clear, meaningful, and relevant. All constructs were measured using seven-point scales. A complete listing of the dynamic capability scale used in the study is provided in appendix D. Reliability analysis indicated that the items for these measures have a Cronbach alpha of 0.809. This conforms to the accepted level of at least 0.70 (Nunnally, 1978). We created the dynamic capabilities index as a linear sum of the dynamic capability items means. The main differences with scales of Danneels (2008) and Drnevich and Kriauciunas (2012) are that the scale is more applicable for new ventures and covers more the different components of a dynamic capability as defined by Teece (2007).

D. MODERATING VARIABLES

Role formalization (RoleForm). Pugh et al. (1963) identified the formalization of organizational tasks and roles as a key attribute of modern organizational structure. Role formalization in entrepreneurial teams captures "what one is asked to do" and refers to the identification and designation of particular functional roles and their assignment to specific

individuals (Dalton et al., 1980). The role formalization variable was adopted from Sine et al. (2006) and is the number of formalized functions in a new venture divided by the potential maximum number of functional roles. The potential functional areas were defined based on Sine et al. (2006) and the face-to-face interviews in the first interview round. These include chief executive officer, chief financial officer, chief engineering/operations officer, human resources, international sales, marketing, research and development, sales, legal/IP. Following Sine et al. (2006), we also orthogonalized the variable role formalization to avoid problems with multicollinearity.

External board (Board). Firms were coded "1" if they have installed an external board. Firms were coded "0" when they did not have an external board. The board can be seen as an external extension of the internal structure and is considered as an important governance mechanism for firm survival (Dalton et al, 1999; Dowell and Schackell, 2011). We only take outside board members into account which means that boards with solely founders and/or members of the management team are not included here.

E. CONTROL VARIABLES

We controlled for company age because this variable could have an important impact on the survival of companies (Sapienza, Autio, George, & Zahra, 2006) and the development of dynamic capabilities (Zahra et al., 2006). Company age is measured in months and collected by using BELFIRST. We use the natural log transformation because the variable company age was skewed. Secondly, we controlled for the environment in which new ventures operate. We used industry-level objective information to derive an index of environmental dynamism. The approach used has been adopted in a number of studies (e.g., Dess and Beard, 1984; Simerly and Mingfang, 2000, Castrogiovanni, 2002) and is viewed as the appropriate level of analysis for studying phenomena related to the environment. The industry-level rate of unpredicted change was measured as the standard errors of two regression slopes following the work of Dess and Beard (1984) and Castrogiovanni (2002). In each case, the independent variable was time. The dependent variables were industry revenues and number of industry employees. Industry revenue has been used as a measure of uncertainty in prior studies (e.g., Keats and Hitt, 1988), and number of employees is a common measure of change in research involving new businesses.

Specifically we regressed industry revenues and industry employees over 5 years against time (2005–2010), and used the standard error of the regression coefficient related to a time dummy variable divided by the average value of industry's revenues and industry employees to produce a standardized index of environmental dynamism. The industry-level archival-based data captured common environmental characteristics faced by participants within a given industry (Boyd, Dess, and Rasheed, 1993). Data on industry revenues and industry employment totals were acquired from the OECD STAN database. Time was regressed

against these variables for the most recent 8-year period. An index of the standard errors of the regression slopes divided by their respective means was used as the indicator of unpredicted change for each of the two variables. These figures were then standardized and summed to create an overall index of environmental dynamism. In addition to this objective measure of environmental dynamism, we also collected perceptual measures of environmental uncertainty and munificence (Maestro, 2009). Maestro (2009) adapted a five-item scale of environmental uncertainty from Miller and Droge (1986) and Sutcliffe (1994). Items include 'Products or services quickly become obsolete in our industry' and 'Actions of competitors in our industry are quite easy to predict (reversed code).' The six-item scale of environmental munificence (Maestro, 2009) was based on Sutcliffe (1994) and Zahra (1993). Items include 'Resources needed for growth and expansion are in abundance and easily accessible in our industry (reverse code)' and 'Demand for products and services in our industry is growing and will continue to grow.'

Beside environmental dynamism and company age, we also controlled for other variables that might influence the impact of dynamic capabilities on new venture survival, such as industry sector. Here, we obtained five categories: ICT, business services, biotech & pharmaceuticals, engineering and manufacturing. As mentioned before, the new ventures in our sample are active in nascent markets which can be brought under this traditional sector classification. Finally, we also controlled for the size of the firm and the founding team. Firm size was the total number of organizational members, including executives and employees. We use the natural log transformation because the variable company size was skewed. Founding team size was the number of executives in a firm. To avoid problems with multicollinearity, founding team size was orthogonalized (Sine et al., 2006). Finally, we also controlled for human resource slack. (Mishina, Pollock, & Porac 2004; Voss, Sirdeshmukh, & Voss, 2008) as this is considered as an important source of dynamic capability (Danneels, 2008). Human resource slack refers to specialized and skilled human resources that are rare and absorbed (Mishina et al., 2004). We measured human resource slack in line with previous recently published works by dividing the number of employees by the total number founding team members (Voss et al., 2008).

IV. FINDINGS

The means, standard deviations and correlations of the variables are presented in Table 1. We found that 32 new ventures (26%) failed to survive the early stages, which is in line with previous studies on new ventures in nascent markets. Roberts (1991) studied technology-based firms in the Boston (US) area and found that failures rates were between 15 and 30%. A Norwegian study showed that survival rate for new technology-based firms is around 75% (Aspelund, Berg-Utby, and Skjevdal, 2005). The companies in our sample are between

6 months and 6 years old and on average 3 years old. Our environmental dynamism measures exhibits similar results as in previous studies (Simerly and Mingfang, 2000).

Table 1. Means, standard deviations, reliabilities, and intercorrelations of study and control variables

	M ean	Std Dev	1	2	3	4	5	6	7	8	9
1. Failure	0.21	0.41	1.00								
2. Csize[b]	4.85	5.22	-0.18*	1.00							
3. Age	34.85	16.12	0.09	0.19*	1.00						
4. Envir	0.01	0.00	-0.17*	0.13	-0.07	1.00					
5. TeamSize[a]	2.62	1.44	-0.18*	0.28**	0.06	0.18*	1.00				
6. DC	5.29	0.69	-0.11	-0.14	-0.15*	-0.02	0.10	1.00			
7. RoleForm[a]	0.29	0.16	-0.33**	0.42**	0.19*	0.14	0.45**	0.04	1.00		
8. Board	0.56	0.50	-0.21**	0.25**	-0.04	0.12	0.37**	0.12	0.44**	1.00	
9. Redundancy[b]	2.61	2.67	-0.01	0.44**	0.19*	0.03	-0.06	-0.17*	0.08	-0.03	1.00

* Significant at $p < 0.05$ (two-tailed), ** Significant at $p < 0.01$ (two-tailed), $n = 170$
[a] Orthogonalized variable
[b] Log-transformed variable

Table 2 summarizes the results of the regression analyses. Because the dependent variable displays the probability of a focal event (firm failure), we employ event history analysis to investigate the impact of dynamic capabilities on firm failure. Event history analysis allows for the modeling of event probability at each time point, and considers both the occurrence and timing of an event, that is, distinguishing between failure one year after company foundation and failure two years after foundation, which is not possible in a logistic regression (Cui et al., 2010). More specifically, we apply a Cox proportional hazard model. Cox models are more suitable than parametric models because it is difficult to make a realistic assumption of the baseline hazard function and incorrect parametric specification of the baseline hazard function would introduce bias into the analysis.

Table 2. Co_x proportional hazard regression with failure as dependent variable

Variables	Model 1	Model 2	Model 3
Sector dummies			
ICT	-0.24 (0.38)	-0.11 (0.33)	-0.08 (0.31)
Business services	-0.14 (0.06)**	-0.62 (0.18)**	-0.32 (0.11)**
Biotech & Pharmaceuticals	-1.23 (0.15)**	-0.52 (0.11)**	-0.51 (0.19)**
Engineering	0.34 (0.10)**	0.56 (0.18)**	0.52 (0.19)**
Founding team size (TeamSize)	-0.22 (0.23)	0.31 (0.30)	0.38 (0.24)*
Company size (Csize)	-0.48 (0.14)**	-0.48 (0.24)*	-0.74 (0.45)*
Company age (Age)	0.01 (0.01)*	0.01 (0.01)†	0.01 (0.01)**
Environmental dynamism (Envir)	-0.15 (0.13)	-0.19 (0.13)	-0.16 (0.10)†
Redundancy	0.15 (0.14)	0.20 (0.22)	0.41 (0.45)*
Role formalization (RoleForm)		-0.70 (0.31)*	-1.04 (0.26)**
External board (Board)		-0.13 (0.30)	0.19 (0.28)
H1 Dynamic Capabilities (DC)		-0.56 (0.18)**	-0.92 (0.16)**
H2 Dynamic Capabilities X Role formalization			-1.04 (0.16)**
H3 Dynamic Capabilities X External board			0.82 (0.49)*
Log-likelihood	-148.57	-139.28	-135.72
Generalized R^2	0.21	0.30	0.38
n	170	170	170

Significance tests are one-tailed for hypothesized relations and two-tailed for controls.
†$p < 0.10$, *$p < 0.05$, **$p < 0.01$

All models in Table 2 are highly significant. Column 1 presents the results of the baseline model with control variables alone ($X^2 = 1501.46$, $p < 0.001$). We see that company age has a positive significant impact on new venture failure, while company size is negatively related to failure. Firms founded in biotech & pharmaceutical emerging markets or firm established around innovative business services are more likely to survive the early stages. Instead, firm active in emerging markets related to the engineering industry are more likely to cease their activities. Environmental dynamism has a negative impact on new venture survival. We also run the same model with perceived environmental uncertainty and found a similar negative relationship with survival. Column 2 shows the results of the regression with the direct effects ($X^2 = 25.32$, $p < 0.001$). After introducing the direct effects in Model 2, the explanatory power, measured by the generalized R square, increases significantly from 0.21 to 0.30. H1 which posits that the dynamic capabilities of a new venture will help new ventures to survive the early stages and thus negatively impacts failure is supported ($p < 0.01$). Further, we noticed that role formalization has a negative significant impact on new venture survival which is in line with the findings of Sine et al. (2006). The final model includes the interaction effects ($X^2 = 14.58$, $p < 0.01$). The generalized R-square increases again significantly from 0.30 to 0.38. H2 which states that role formalization has a moderating effect on the relationship between dynamic capabilities and failure is supported at the 0.01 level. H3, which stated that the negative impact of dynamic capabilities on the probability of failure will decrease with the availability of an external board, also found support ($p < 0.05$). The simple slope analyses (illustrated in figure 1and 2) confirm our interpretations of the moderating effects in the regression analyses.

Figure 1. Graphical Presentation of Interaction between 'dynamic capabilities' and 'Role formalization' on new venture failure

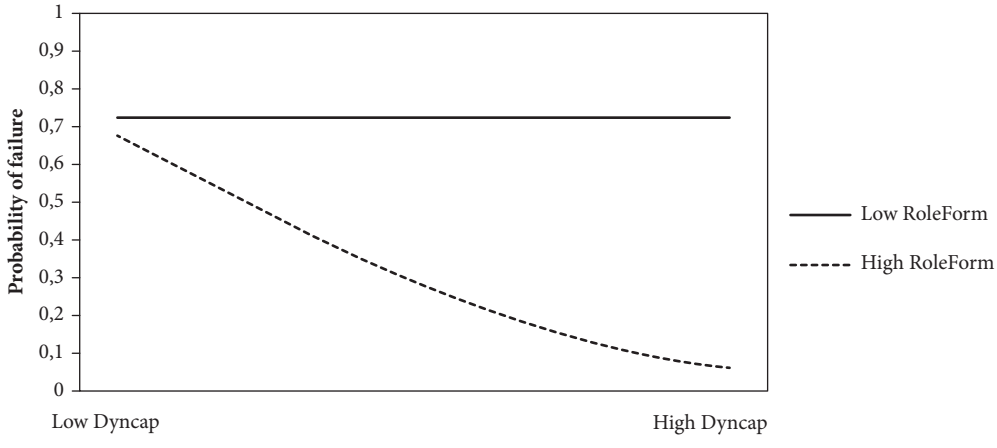
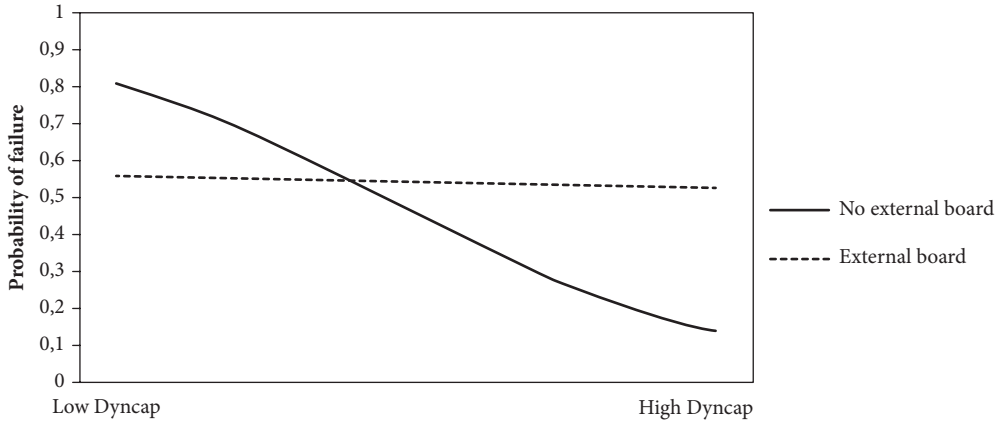


Figure 2. Graphical Presentation of Interaction between 'dynamic capabilities' and 'External Board' on new venture failure



V. DISCUSSION AND CONCLUSIONS

Our objective in this study was to examine the relationship between the dynamic capabilities developed in a new venture, the mechanisms which invoke its stability and its survival in nascent markets. Combining arguments based upon organizational theory and dynamic capabilities, we show that the internal structure of the venture in terms of the adopted role formalization in the founding team form a strong boundary condition for dynamic capabilities to enhance the survival potential of the venture. As expected, developing dynamic capabilities over time also increases the potential survival of these firms significantly. Finally, boards do not amplify the impact of dynamic capabilities on survival. On the contrary, they moderate this relation.

This study extends the literature on dynamic capabilities by showing how stability in its form of formal structures within the organization forms fertile grounds. In line with Farjoun's (2010) theoretical arguments on duality between change and stability, we find that dynamic capabilities work best when they are embedded in a stable and well-structured organization. As such, they do not form a tradeoff but have, as hinted at by Zahra et al. (2006), complex interactions among each other. Dynamic capabilities without stable underlying structures might lead to chaos and at least moderate the positive impact of these capabilities in nascent markets. The counter-intuitive hypothesis that dynamic capabilities are strengthened by this form of stability is novel and extends the theoretical perspective on dynamic capabilities, which at most considers these capabilities to be contingent upon the environment and the development of underlying operational capabilities. The fact that organizational stability underpins the impact of dynamic capabilities has been largely

neglected in the literature. Our study shows that companies in nascent markets, of which the environmental conditions force these companies to explore business models, need to create stability in order to be able to facilitate changes by structuring and bundling key resources in substantive capabilities (Sirmon et al., 2007). These findings are also in line with Sine et al.'s (2006) findings that formal structure enhances performance in new ventures (in nascent markets). However, we show that their findings do not exclude flexibility. Ventures which have a degree of formalization exceed in addressing changes in the environment by developing specific dynamic capabilities, which in turn amplify the impact on performance.

In addition to the theoretical contribution, this article also provides an empirical contribution to the literature on dynamic capabilities (Eisenhardt and Martin, 2000; Teece, 2007) by developing and testing measures of dynamic capabilities. Moreover, we show the often assumed but never tested positive impact of dynamic capabilities on the most important performance outcome of new ventures in nascent markets, namely firm survival. Despite the increased interest in dynamic capabilities, the concept remains an empirically unexplored construct with the exception of a few studies (eg. Danneels, 2008; Drnevich and Kriauciunas, 2012), most researchers measure the construct in an indirect way instead of developing a scale. The fact that we empirically show that dynamic capabilities contribute to the survival of ventures in nascent markets, reinforces the underlying assumption that developing the conditions in a ventures which allow for change, prevails.

Third, our paper also shows that despite the fact that the board literature indicate that boards bring stability to a venture due to their monitoring function, they do not amplify the impact of dynamic capabilities. We build on the notion of resource cognition (Danneels, 2010) to show that the further away executive decision makers are from the work floor the more difficult it will be for them to assess the real resources of the company. Hence, their deep involvement in the strategic decision making might be counterproductive, especially in nascent markets where usually no financial performance indicators are possible to use as benchmarks. This finding is particularly important to increase our understanding of the role which boards play in new ventures and extends the relatively new, emerging theories on board monitoring in new ventures (Garg, 2012).

REFERENCES

Aspelund, A., Berg-Utby, T., & Skjevdal, R. 2005. Initial resources' influence on new venture survival: a longitudinal study of new technology-based firms. *Technovation*, 25: 1337–1347

Arora, A., & Nandkumar, A. 2011. Cash-Out or Flameout! Opportunity Cost and Entrepreneurial Strategy: Theory, and Evidence from the Information Security Industry. *Management Science*. 57(10): 1844–1860.

Arthurs, J.D., & Busenitz, W. 2006. Dynamic Capabilities and Venture Performance: The Effects of Venture Capitalists, *Journal of Business Venturing*, 21: 195–215.

Autio, E., George, G., & Alexy, O. 2011. International Entrepreneurship and Capability Development-Qualitative Evidence and Future Research Directions. *Entrepreneurship theory and practice,* 35(1), 11–37.

Barreto, I.K. 2010. Dynamic Capabilities: A Review of Past Research and an Agenda for the Future, *Journal of Management,* 36: 256–278.

Bingham, C., Eisenhardt, K., & Davis, J. 2009. *Opening the Black Box: What firms explicitly learn from their process experience,* Working paper, University of North Carolina.

Bingham, C.B. 2009. Oscillating improvisation: How entrepreneurial firms create success in foreign market entries over time. *Strategic Entrepreneurship Journal*, *3*(4), 321–345.

Boyd, B.K., Dess, G.G., & Rasheed, A.M. 1993. Divergence between archival and perceptual measures of the environment: Causes and consequences, *Academy of Management Review,* 18: 204–226.

Castrogiovanni, G.J. 2002. Organization task environments: have they changed fundamentally over time. *Journal of Management,* 28(2): 129–150.

Certo S.T., Daily, C.M., & Dalton, D.R. 2001. Signaling firm value through board structure: an investigation of initial public offerings. *Entrepreneurship Theory and Practice,* 26(2): 33–50.

Cui, A.S., Calantone, R.J, & Griffith, D.A. 2010. Strategic change and termination of interfirm partnerships. *Strategic Management Journal*, 32: 402–423.

Dalton, D.R., Todor, W.D., Spendolini, M.J., Fielding, G.J. & Porter, L. 1980. Organization Structure and Performance: A Critical Review. *Academy of Management Review*, 5: 49–64.

Danneels, E. 2008. 'Organisational Antecedents of Second-Order Competences', *Strategic Management Journal*, 29: 519–543.

Danneels, E. 2010. 'Trying to become a different type of company: dynamic capability at Smith Corona', *Strategic Management Journal*, 32: 1–31.

D'Aveni, R.A., Dagnino, G.B., & Smith, K.G. 2010 The age of temporary advantage. *Strategic Management Journal*.31: 1371–1385.Dess GG, Beard DW. 1984. Dimensions of organizational task environments. *Administrative Science Quarterly,* 29: 52–73.

Dowell G., & Schackell, M. 2011. Boards, Ceos, and surviving a financial crisis: evidence from the internet shakeout. *Strategic Management Journal*, 32(10): 1025–1045.

Drnevich P.L., & Kriauciunas, A.P., 2011. Clarifying the conditions and limits of the contributions of ordinary and dynamic capabilities to relative firm performance. *Strategic Management Journal,* 32: 254–279

Eisenhardt, K.M., & Martin, J.A. 2000. Dynamic capabilities: what are they?' *Strategic Management Journal,* 21: 1105–1121.

Farjoun, M. 2010. Beyond Dualism: Stability and Change as a Duality, *Academy of Management Review*, 35(2): 202–225.

Garg, S. 2013. Venture boards: Differences with public firm boards and implications for monitoring and firm performance, *Academy of Management Review,* 38: 90–108.

Gruber, M., Heinemann, F., Brettel, M., & Hungeling, S. 2010. Configurations of resources and capabilities and their performance implications: an exploratory study on technology ventures. *Strategic Management Journal,* 31(12): 1337–1356.

Hannan, M., & Freeman, J. 1989. *Organizational Ecology*. Harvard University Press: Cambridge, MA.

Helfat, C.E., & Peteraf, M.A. 2003. The dynamic resource-based view: Capability Lifecycles. *Strategic Management Journal,* 24: 997–1010.

Helfat, C.E., & Peteraf, M.A. 2009. Understanding dynamic capabilities: progress along a developmental path. *Strategic Organisation,* 7(1): 91–102.

Helfat, C.E., & Winter, S.G., 2011.Untangling dynamic and operational capabilities: strategy for the (n)ever-changing world. *Strategic Management Journal,* 32(11): 1243–1250.

Kaplan, S., & Tripsas, M. 2008. Thinking about Technology: Applying a Cognitive Lens to Technical Change, *Research Policy*, 37: 790–805.

Keats, B.W., & Hitt M.A. 1988. A causal model of linkages among environmental dimensions, macro organizational characteristics, and performance, *Academy of Management Journal*, 31:570–598.

Maestro M. 2009. *Polychronicity in top management teams: Implications for group processes, strategic decision making & new venture performance*. Ph.D Imperial College London.

Makadok, R. 2001. Toward a synthesis of the resource-based and dynamic-capability views of rent creation. *Strategic Management Journal*, 22: 387–401.

Mishina, Y., Pollock, T. G., & Porac, J. F. 2004. Are more resources always better for growth? Resource stickiness in market and product expansion. *Strategic Management Journal*, 25: 1179–1197.

Miller, D., & Droge, C. 1986. Psychological and traditional determinants of structure, *Administrative Science Quarterly,* 31(4): 539–560.

Nunnally, J.C. (1978). *Psychometric theory (2nd ed.).* New York: McGraw-Hill.

Ozcan, P., & Eisenhardt, K. 2009. Origin of Alliance Portfolios: entrepreneurs, network strategies and firm performance, *Academy of Management Journal*, 2:246–279.

Pugh, D.S., Hickson, D.J., Hinings, C.R., Macdonald, K.M., Turner, C., & Lupton, T. 1963. A conceptual scheme for organizational analysis. *Administrative Science Quarterly*, 8: 289–315.

Santos, F., & Eisenhardt, K. 2009. Constructing Markts and Shaping Boundaries: Entrepreneurial Power in Nascent Fields, *Academy of Management Journal*, 52(4): 643–671.

Sapienza, H.J., Autio, E., George, G., & Zahra, S.A. 2006. A capability perspective on the effects of early internationalization on firm survival and growth. *Academy of Management Review*, 31: 914–933.

Schreyögg, G., & Sydow, J., 2010. Organizing for fluidity? Dilemmas of new organizational forms. *Organization science*, 21(6): 1251–1262.

Simerly, R., & Mingfang, L., 2000. Environmental dynamism, capital structure and performance: a theoretical integration and an empirical test. *Strategic Management Journal* 21: 31–49.

Sine, W.D., & David, R. 2003. Environmental Jolts, Institutional Change and the Creation of Entrepreneurial Opportunity in the US Electric Power Industry, *Research Policy*, 32: 185–2007.

Sine, W.D., Mitsuhashi, H. & Kirsch, D.A. 2006. Revisiting Burns and Stalker: Formal Structure and New Venture Performance in Emerging Economic Sectors, *Academy of Management Journal*, 49(1): 121–132.

Sinha, R.K, & Noble, C.H, 2008. The adoption of radical manufacturing technologies and firm survival. *Strategic Management Journal*, 29: 943–962.

Sirmon, D. G., Hitt, M. A. & Ireland, R. D. 2007. Managing firm resources in dynamic environments to create value: Looking inside the black box. *Academy of Management Review*, 32: 273–292.

Sirmon, D., Hitt, M.A., Ireland, R.D., & Gilbert, B.A. 2011. Resource orchestration to create competitive advantage: Breadth, depth, and life cycle effects. *Journal of Management,* 37: 1390–1412.

Smith, W.K., & Lewis, M.W. 2011. Towards a theory of paradox: a dynamic equilibrium model of organizing. *Academy of Management Review*, 36(2): 381–403.

Stinchcombe, A. 1965. Social structure and organizations. In J. March (Ed.), *Handbook of organizations:* 142–193. Chicago: Rand McNally.

Sutcliffe, K.M. 1994. What executives notice: Accurate perceptions of top management teams. *Academy of Management Journal,* 37: 1360–1378.

Teece, D.J., Pisano, G., & Shuen, A. 1997. Dynamic capabilities and strategic management. *Strategic Management Journal,* 18: 509–533.

Teece, D.J. 2007. Explicating dynamic capabilities: The nature and microfoundations of (sustainable) enterprise performance. *Strategic Management Journal,* 28: 1319–1350.

Tuggle, C., Sirmon, D., Reutzel, C. & Bierman, L. Commanding board of director attention: Investigating how organizational performance and ceo duality affect board members' attention to monitoring. *Strategic Management Journal,* 31: 946–968.

Voss, G.B., Sirdeshmukh, D., & VOSS, Z.G. 2008. The effects of slack resources and environmental threat on product exploration and exploitation. *Academy of Management Journal.* 51(1):147–164.

Wasserman, N. 2006. Stewards, agents and the founder discount: executive compensation in new ventures, *Academy of Management Journal*, 49:960–979.

Wasserman, N., & Boeker, W. 2010. *Mentoring and Monitoring: Boards of Directors and CEO in New Ventures.* HBS Working Paper, Camebridge, MA.

Winter, S. 2003. Understanding dynamic capabilities, *Strategic Management Journal* 24: 991–995.Zahra S. 1993. Environment, corporate entrepreneurship, and financial performance: A taxonomic approach. *Journal of Business Venturing,* 8: 319–340.

Zahra, S., & Nielsen, A.P. 2002. 'Sources of capabilities, integration and technology commercialization'. *Strategic Management Journal*, 23, 377–98.

Zahra, S., Hitt, M.A., & Ireland R.D. 2000. International Expansion by New Venture Firms: International Diversity, Mode of Market Entry, Technological Learning, and Performance. *Academy of Management Journal,* 43(5), 925–950.

Zahra, S., Sapienza, H., & Davidsson, P. 2006. Entrepreneurship and dynamic capabilities: a review, model and research agenda. *Journal of Management Studies*, 43: 917–955.

Zollo, M., & Winter S.G. 2002. Deliberate learning and the evolution of dynamic capabilities. *Organization Science,* 13: 339–351.

APPENDIX A: DYNAMIC CAPABILITIES SCALE

Different companies are good at different things. The following questions ask you to assess your company's skills in various areas, relative to other start-ups. Relative to other start-ups and based on new information, my company is good at …

	Strongly Disagree Strongly Agree						
1. Assessing the potential of new markets	1	2	3	4	5	6	7
2. Researching new competitors and new customers	1	2	3	4	5	6	7
3. Assessing the feasibility of new technologies	1	2	3	4	5	6	7
4. Identifying promising new technologies	1	2	3	4	5	6	7
5. Changing the marketing and communication plan	1	2	3	4	5	6	7
6. Altering the product roadmap	1	2	3	4	5	6	7
7. Making changes to the global delivery model / distribution channels	1	2	3	4	5	6	7
8. Revising the technology roadmap	1	2	3	4	5	6	7

THE IMPACT OF DIFFERENT BODY-SIZES OF NON-CELEBRITY ENDORSERS ON ADVERTISING EFFECTIVENESS

Irene Roozen*

Abstract

Many studies have shown that attractive endorsers are more successful in creating positive attitudes towards advertisements than their less attractive counterparts. This paper focuses on the attractiveness of the body-sizes of the endorsers. In a first study 'ideal' body-sizes of female and male endorsers are investigated in an experiment with different photo-shopped endorsers. The results of these findings are used in two further experiments in which the influence of endorsers' body-sizes on the effectiveness of print advertisements are investigated. Generally, endorsers with ideal body-sizes are the most effective. However, the results show that this 'general' rule does not always hold and that personal characteristics of the previewer, notably his/her body esteem, gender and the sex of the audience have an important significant impact on the advertising effectiveness of the 'ideal' body-size endorsers.

Keywords: advertising; body esteem; endorsement; gender differences; self-esteem

JEL codes: M31, M37

I. INTRODUCTION

Showing idealized thin female models as endorsers of products in advertisements is a common practice in advertising. However, at the same time, it has also elicited a lot of criticism by different actors in society, such as doctors, psychologists, public policymakers, journalists and academic researchers. Academic research shows that some girls and women, when exposed to such idealized images, could start to feel negative about themselves, which

* Irene Roozen, Human Relations Research Group, FEB, KU Leuven Campus Brussel, Warmoesberg 26, 1000 Brussel. Email: irene.roozen@kuleuven.be.

could further lead to severe health problems, such as depression and eating disorders. (Halliwell, Dittmar & Howe, 2005; Bessenoff, 2006).

Despite these negative reactions, advertisers continue the use of idealized female models in their campaigns. Editing photographic images of models to make them look better is a commonly used technique. In October 2009, the model Filippa Hamilton of Ralph Lauren was fired, because she was considered to be too fat for a model (she weighed 54 kg and was 1 meter 78 tall at that moment). She took court action and the American fashion house admitted that they had been photo editing her image in the advertisements (the image showed her head being larger than her waist) (see the story on www.youtube.com/watch?v=DQ9p0 MxqQAQ).

However, there are also advertisers who react against using stereotypically thin female images of beauty in the media and follow their own approach, such as Dove with their "Real Beauty" campaign in the period 2003–2007. This Dove campaign won several awards which honor significant achievement in marketing communications effectiveness (e.g. Grand Prix Cannes Advertising Awards in 2007, the silver IPA for effectiveness and a Grand EFFIE, (Marketingheart.wordpress.com, 2011; Bissell & Rask, 2010) and received a lot of additional positive attention from consumers and in the media (on television, for example, in the Oprah Winfrey Show, The Ellen DeGeneres Show). However, the Dove campaign is still an exception and idealized thin female models are still preponderant in advertisements. The question, therefore, remains whether the use of thin female models is always significantly more effective.

II. AN OVERVIEW OF THE LITERATURE

In the literature on endorsement, many studies are focused on *celebrity* endorsement and show that physically attractive celebrities are more effective in an advertisement (Bower, 2001; Milkie, 1999, 2002, Smeesters et al. 2010). Advertising effectiveness is defined in the literature in terms of improving the attitude towards the advertisement, the attitude towards the product and the intention to purchase the product (Roozen & Claeys, 2010; Erdogan et al., 2001). The endorsers are significantly more successful in creating positive attitudes towards the advertisements and towards the products and or brands used in the advertising (Erdogan, 1999; Eisend & Langner, 2010). However, there is no conclusive evidence that attractive endorsers are always able to create significantly more purchase intentions (Erdogan et al., 2001).

For celebrities, the meaning-transfer model of McCracken (1989) is often used to emphasize the endorsement effectiveness. According to this model (McCracken, 1989) the process of celebrity endorsement consists of three subsequent stages. First, the positive feelings associated with the famous person are passed on to the product or brand. In the second stage, the positive feelings become associated with the product or brand in the consumer's mind. Finally, the consumer identifies himself with the symbolic properties of

the product. The process of meaning transfer is now completed. According to the meaning-transfer model the endorser with an 'ideal' body-size should yield the highest positive feelings and finally the highest scores on the advertisement effectiveness. The product match-up model is based on the idea that the celebrity's image and the product message should be congruent (Bower & Landreth, 2001; Erdogan, 1999). The balance theory combines the principles of the source-attractiveness model and the match-up hypothesis. According to this theoretical framework, celebrity endorsers can serve as a marketing tool when two conditions are met: First, the endorser is well-liked by the consumer and second, there is a match between the celebrity and the endorsed brand or product. When both conditions are fulfilled, consumers transfer their positive feelings about the celebrity endorser to the product and are more likely to buy it (Solomon, Bamossy, Askegaard, & Hogg, 2006). According to this theory, models with an 'ideal' body-size endorsing products which are related to the attractiveness of body-sizes should be more effective in advertising than their counterparts.

Academic research points to inconsistencies in the effectiveness of attractive models in advertising campaigns and suggests different moderators, mostly product-related (cfr. the match-up hypothesis, Bower & Landreth, 2001), but also here, results have not always been conclusive (Kang & Herr, 2006; Till & Busler, 2000). Moreover, attractiveness of an endorser is a broad concept and the literature shows different interpretations of 'attractive' endorsers. The physical appearance of the endorser (i.e. clothes, colour of the eyes, hair) is often taken into account (Erdogan, 1999; Kilbourne, 1990) but also lifestyle, intellectual capabilities, athletic/sport performances are used to indicate the 'attractiveness' of an endorser, however, the body-size of an endorser as a component of 'attractiveness' has not being studied. Furthermore, besides product-related moderators, individual differences, such as socio-demographic and psychographic variables, could also be important. Considering prior research on the effects on levels of women's body- and self-esteem of attractive models (Grabe et al., 2008; Groesz, Levine & Murnen, 2002; Polivy & Herman, 2002; Henderson-King & Henderson-King, 1997; Irving, 1990; Joshi, Herman & Polivy, 2004; Solomon et al. 2008; Halliwell, Dittmar, 2004; Halliwell, Dittmar & Howe, 2005; Bessenoff, 2006; Jalees & Majid, 2009), one could also expect a significant moderating effect of people's chronic level of esteem (Higgins & Brendl, 1995). That is, based on people's chronic level of body- and self-esteem, they could be differently influenced by external primes and specifically, by different body-sizes of (female) endorsers (Mendelson, et al. 2002; Mathes & Kahn, 1975; Martin et al. 2007; Harrison & Cantor, 1997). For example, a meta-analytic review of Groesz et al. (2002) showed that based on 25 studies, the body image of females was significantly more negative after viewing thin models (especially for younger female participants) than after viewing images of either average body-size or 'plus' body-size models (a significant effect size was found $d = -.31$). However, to our knowledge, such moderators have been somewhat neglected in prior experimental research on advertising effectiveness.

Despite some notable exceptions (e.g., Dens, De Pelsmacker & Janssens 2009), most prior studies have focused on female celebrity endorsers and female audiences (e.g., Bower, 2001; Bower & Landreth, 2001; Jalees & Majid, 2009), while in practice men are becoming a more and more important target group for advertisers. While women do most of the shopping, men increased their shares in shopping trips between 2004 and 2011 in all retail channels (Mahoney, 2011). Also, more products are uniquely developed for men, such as beverages, beauty and hygiene products, which is clear from the use of more male models in advertising campaigns (Costa, 2011). Although research on male body image has increased, it is still quite limited in scope compared to female models (Bottamini & Ste-Marie, 2006; Agliata & Tantleff-Dunn, 2004; Diedrichs & Lee, 2010; Smolak et al. 2005). In particular, media-portrayed male endorsers have hardly been researched into for advertising effectiveness. However, Grogan (2008) suggests that most men aspire to a muscular mesomorphic shape characterized by well-developed muscles on chest, arms, and shoulders, rather than a slim or fat build. Bottamini and Ste-Marie (2006) also found in their qualitative research that the majority of the men expressed a desire to develop muscles to obtain their desired physique. It remains to be seen whether male and female audiences consider that these characteristics define an 'ideal' body size for men.

In conclusion, while the literature contains many studies that have looked at the role of body sizes in advertisements, there remain important gaps in our knowledge of the impact of using different body sizes of endorsers on the advertisement effectiveness.

Firstly, the literature shows divergent results when it comes to the definition of 'ideal' body sizes which indicates that the definition of 'ideal' body-sizes of endorsers remains uncertain.

Secondly, research is often based on celebrity endorsers and relatively little is known about the 'ideal' body-size of male and female non-celebrity endorsers.

Thirdly, the literature review indicates that studies on the effectiveness of endorsers in advertisements do not allow to filter out the contribution that ideal female and male body sizes make to the effectiveness of the advertisement since comparisons of advertisements with different body sizes do not correct for other characteristics of the endorser or the advertisement (e.g. celebrity status, body positioning, advertising context). There is no research to our knowledge that allows to assess the contribution of different male and female body-sizes of endorsers to the effectiveness of the advertisement.

Fourthly, the literature does not assess whether the ad effectiveness of ideal body sizes of endorsers is affected by moderating factors (psychographic and socio-demographic characteristics of the viewers and product characteristics) – even if there is a lot of research on the effects of endorsers' body sizes on self-esteem and body esteem of female viewers.

Finally, prior research often relies on specific audiences, e.g. female respondents with specific characteristics (e.g. eating disorders), which makes it difficult to draw more general conclusions.

III. PROBLEM STATEMENT AND RESEARCH DESIGN

Against this background, this paper attempts to answer the following three more generic research questions:

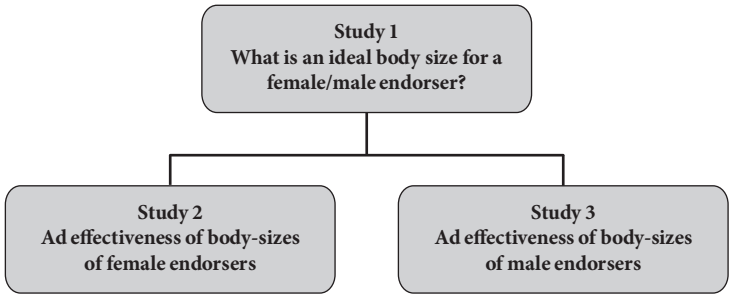
RQ 1. What is an 'ideal' body-size for an unknown (i.e. non celebrity) female/male endorser?

RQ 2. Are 'ideal' female/male body-size endorsers more effective in print advertisements? Do 'ideal' body-size endorsers have a significantly higher score on the attitude towards the advertisement, the attitude towards the product, purchase intention of the product, and the attitude towards the endorser (the inner- and outer characteristics of the endorser)?

RQ 3. Are there moderating factors (socio-demographic characteristics (e.g. age and gender) and psychographic characteristics (e.g. chronic levels of self-esteem and body esteem) of the viewer and or product characteristics of the endorsed products in the advertisement), which might influence the advertising effectiveness?

To test the above research questions, we investigate in Study 1 the 'ideal' body-size of unknown female and male endorsers. In Study 2 and 3, we set up two experiments to examine the advertising effectiveness of using different body-sizes of unknown female and male endorsers. We analyse the effectiveness of the 'ideal' body-sizes and, in addition, investigate if certain individual- and or product characteristics significantly influence the effectiveness of the body-size of the endorser in the print advertisement (research questions two and three). In Figure 1 an overview of the three studies is presented.

Figure 1. Overview of the studies



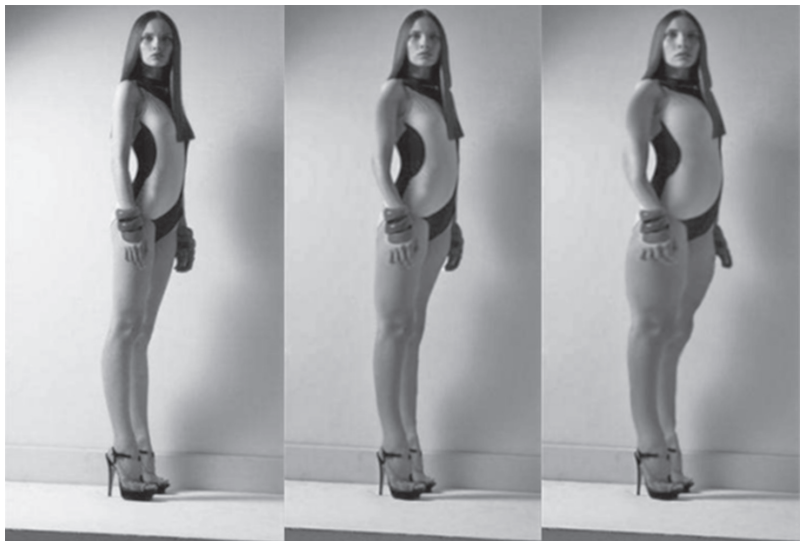
The remainder of this paper is organized as follows. The next part sets out the method, design, sample and results of the experiments. In the first study the focus is on the ideal body-size for female and male endorsers. In the second study the impact of the body-shape of female images in print advertisements is analysed and in the third study the same experiment using

male images is described. Finally, part four compares the experiments and discusses some tentative implications.

IV. STUDY 1

For the first study, we chose two unknown (non-celebrity) female endorsers and two unknown male endorsers who differed in their general looks in order to prevent confounding effects as much as possible. The two different female models were given three different body-sizes (i.e., relatively slim, average size and full-figured). In order to realize this, we used editing software (Photoshop) to manipulate the size while leaving the 'look' of the face unaltered (for an example, see Figure 2). The relatively slim model was designed according to the measurements of the average female fashion model as listed on the agency websites (size small 'relatively slim model' = (fr) 34/36; www.models.com). Based on the literature, we believe that the relative slim model matches best with an 'ideal' model. We estimated the average body-size model at size (fr) 38–40 and the full-figured body-size model at size (fr) 42–44.

Figure 2. Example of three different female models (left: slim (ideal) model, average-size, full-figured model)



Also the two different male models were given three different body-sizes. As is discussed in the literature review, most men aspire to a muscular mesomorphic shape, which is characterized by well-developed muscles on chest, arms and shoulders (Grogan, 2008). Grogan (2008) indicates that men who are dissatisfied with their body shape would like to

have more muscles (e.g. biceps, shoulders, chest) which can actually mean both an increase or a reduction in body weight, unlike 'dissatisfied' women who mostly want to become slimmer. In our research, the average measurements of the models were European shirt size = medium, chest = 100 cm, waist = 86 cm and hips = 105 cm. These sizes are comparable to those of the average Belgian male fashion model (as listed on the agency websites, www.models.be). The slim model, in our study, has average European shirt size small (chest = 80 cm; waist = 70 cm; hips = 86 cm) and the full-figured model has average European shirt size large (chest = 112 cm. waist = 98 cm, hips = 116 cm).

In Figure 3, examples of a slim, average and full-figured male endorser are given.

Figure 3. Example of three different male models (left: slim model, average-size, full-figured (ideal) model)

As shown in the examples of Figure 2 and 3, we have opted to analyze 'dressed' (female and male) endorsers. The earlier mentioned male body-size of a muscular mesomorphic shape, characterized by well-developed muscles on chest, arms and shoulders is therefore less pronounced because the clothes cover the torso. In this research we, therefore, investigate an 'ideal' male 'dressed' endorser instead of analyzing an 'ideal' bodybuilder.

A. DESIGN AND PROCEDURE

The study focused on body-sizes of different female and male endorsers. In total 12 endorsers with different body-sizes (four models times three body-sizes) were generated and randomly presented to the previewers. In particular, all respondents saw three times the four different models (female model 1, male model 1, female model 2, male model 2) with randomly three

different body-sizes. The models were unknown (i.e., no celebrities or well-known models and royalty-free Internet pictures).

One of the greatest advantages of this within subjects' design is that it reduces the errors associated with individual differences of the participants because all the respondents are exposed to all the different endorsers which means that each respondent serves as his or her own baseline. In a between-subjects' design, the respondents may differ with regard to important individual characteristics that can have an impact on the dependent variables (e.g. gender, age, self-esteem, body-esteem, eating disorders, social desirability of answering, …). A within subjects' design reduces the error variance because any factor that may influence the dependent variable (in our case: 'score on the ideal body-size') is exactly the same for the different conditions (in our case: the different 'shown' pictures of the endorsers), because the respondents are the same group of people for the different conditions. Another main advantage is that the within subjects' design does not require a large sample. This is why we have opted for a within subjects' design. However, carryover effects (participants are tested several times to the same kind of treatment which can affect their answers) and fatigue (i.e. the total length of the questionnaire increase) could be significant major drawbacks. We have tried to decrease these effects by randomly presenting the 12 different endorsers (taking order effect into account) and asking only one question per 'treatment' endorser e.g. please, rate the score the body-size of the model on the picture on a 1 to 10 scale.

B. SAMPLING METHOD

We ran this study among male and female respondents from different age groups (i.e., young adults (from 18 to 25 years old) and adults (older than 26 years)). The main data collection was carried out via a self-selected online survey of master degree students and their relatives at a university in Belgium. In total, 97 completed responses were used for our data analysis. The majority of the sample was female, 80%. Moreover, 63% of the respondents were between 18 and 25 years old. No significant differences were found between age and gender (X^2 (1) = 2.54, p = .111).

C. RESULTS

A repeated-measures ANOVA, with a body-size of the model as a within-subjects factor, revealed that there was a significant main effect for body-sizes for all the female and male models (i.e., relatively slim, average size, full-figured). For the female models the endorsers with the relatively small body-size score significantly higher on the scale for ideal body-size than the full-size models (female model 1 (F[2, 191] = 49.649, p < .001; female model 2 (F[2, 192] = 55.659, p < .001)). For both the male models the full-size endorsers score significantly

higher (male model 1 (F[2, 192] = 41.447, p < .001 = .037), male model 2 (F[2, 192] = 131.55, p < .001)). In Table 1 the averages and standard deviations of the 12 endorsers are presented.

Table 1. Average scores (standard deviations) of the endorsers with different body-sizes (n = 97)[*)]

Body-size	Female model 1	Female model 2	Male model 1	Male model 2
Small/Slim	**7,52 (2,04)**	**7,14 (1.99)**	5,89 (1.92)	3,53 (2.04)
Average	7,08 (2,18)	5,42 (1.82)	6,72 (1.72)	6,20 (1.66)
Full-size /muscular	5,29 (1.95)	4,82 (1.63)	**7,18 (1.70)**	**7,15 (1.60)**

[*)] Post-hoc tests show significant differences between the different groups (small, average, full-size/muscular) for female model 1, 2 and male model 2. For male model 1 no significant differences were found between the 'average' and the 'muscular' endorser.

Furthermore, no significant influence was found for gender and the age of the participants and the 'ideal' body-size scores of the endorsers. These findings give an answer on the first research question, a slim female and a muscular male model are considered as 'ideal' body-sizes of endorsers. In Study 2 and 3 we analyse if the body-size of the endorser can significantly influence the ad effectiveness.

V. STUDY 2

A. DESIGN AND PROCEDURE

For Study 2 we selected three unknown, professional, female models who differed in their general looks in order to prevent confounding effects as much as possible. The first female model wears a (black/white) swimming suit, the second a short, black dress and the third an open black leather jacket with black lingerie underneath. These female models were given, three different body-sizes (i.e., relatively slim, average size and full-figured) by using editing software (Photoshop). The three different body-size related products in the advertisement were: body lotion, perfume and a fashion magazine. These products all received fictitious brand names to control for existing brand preferences and knowledge which could possibly distort respondents' evaluation of the advertisements. This set-up generated 27 (= 3 × 3 × 3) different advertisements. However, as a between-subjects design would require an enormous amount of data collection, we decided to set up a within subjects' design. This design also reduces the errors associated with individual differences of the participants. In Figure 4, our research design, a 3 × 3 Latin square mixed factorial design, is given. In particular, all respondents saw all three specific products and models of all three sizes, but in random order. Respondents in the first experimental group were first exposed to a perfume print advertisement with a slim female wearing a black dress,

whereas respondents in the second experimental group evaluated the same ad with an average-sized model and in the third experimental group with a full-figured model (see Figure 4).

Figure 4. Research design: 3 × 3 Latin square mixed factorial design

Print ad	Experimental group I	Experimental group II	Experimental group III
1. Ad-Perfume	Slim model Black dress	Average-size model swimsuit	Full-figured model Leather jacket
2. Ad- magazine	Full-figured model swimsuit	Slim model Leather jacket	Average-size model Black dress
3. Ad-body lotion	Average-size model Leather jacket	Full-figured model Black dress	Slim model swimsuit

The literature indicates that body esteem and self-esteem of the viewer could have important effects on the perception of ideal body sizes (e.g. Groesz, Levine & Murnen, 2002; Halliwell & Dittmar, 2004; Grab, Ward, & Hyde, 2008; Halliwell, Dittmar & Howe, 2005; Posava, Posavac & Posavac, 1998; Hobza, et al. 2007; Lorenzen, Grieve & Thomas, 2004; Dens et al. 2009). This suggests that these characteristics might also have moderating effects on the ad effectiveness of different body sizes. We have, therefore, measured them explicitly (see below).

B. SAMPLE

We ran this study also among male and female respondents from different age groups The main data collection was carried out via a self-selected online survey. In total, 266 completed responses were used for our data analysis. The majority of the sample was female, 61.7%. Moreover, 59% of the respondents were between 18 and 25 years old, 24% fell into the age category 26–45 years old, and 17% were 46 years old or older. In the study, the respondents were randomly assigned to one of the three different experimental groups. No significant differences were found for age and gender between the different experimental groups (X^2_{gender} (2) = 1.05, p = .59; X^2_{age} (4) = 4.90, p = .30).

C. MEASUREMENTS

The first page of the online questionnaire was an introduction. Participants were informed that the study was undertaken by our university and that there were no commercial intentions. We also instructed them to report their own opinion, assuring them that there were no right or wrong answers. On the second page, we instructed them to carefully look at the ad and

answer the questions related to the ad. On the next two pages, they received the two other ads. Directly after the presentation of the ad, the measurement of the effectiveness of the ad takes place by measuring the attitude towards the advertisement, the attitude towards the product / brand shown in the ad and the attitude towards the endorser and finally the intention to purchase the product.

A 6-item-7-point Likert scale to measure the attitude towards the advertisement was used, based on Spears & Singh (2004). Items included: *'I think the above ad is appealing, attractive, informative, believable, convincing and liking'*. Based on the findings of the Cronbach's alpha and the factor analyses (α = .96; total variance explained 82.0%), we decided to construct one concept for measuring the attitude towards the ad (Aad).

A 5-item-7-point Likert scale based on Gelb & Zinkhan (1986) was used to measure the attitude towards the brand 'AB' *('I think the brand shown in the advertisement is good, interesting, high quality, nice, tasteful'*, α = .93; 79.0%). For purchase intentions (PI) a 4-item-7-point Likert scale based on Jamieson (1989) was used *('I would like to try this product, If I could choose I would try this product, I plan to buy this product, I am eager to check out this product'*; α = .89; 75.0%). The attitude towards the endorser is based on the scale of Spears and Singh (2004) to measure the perceived inner and outer characteristics. We first measured the perceived 'inner' characteristics of the endorser ('Att-Endorser-Inside') by asking if the endorser is trustworthy, sincere and honest (α = .95, 90.4%) and afterwards the perceived 'outer' characteristics ('Att-Endorser-Outside') by asking if the respondent finds the endorser beautiful, elegant, high class and sexy (α = .95, 86.0%). We have also asked the (female) respondents to compare themselves with the female endorsers in the advertisements and to indicate the 'fit' between themselves and the endorser. A 3-item-7-point Likert scale based on Bower and Landreth (2001) was used ('*I feel there are many similarities between the model and me*'; '*I can identify myself very well with the model*' and, '*I feel that the endorser in the advertisement looks like me';* α = .94; 90.00%). Finally, for every respondent, self-esteem (based on Rosenberg, 1965; α = .92; 77.0%) and body esteem (a 2-item 7-Likert scale based on the body esteem-appearance scale, based on Mendelson et al. 2002; $r = .68$), gender and age were measured. The average time to fill out the questionnaire was 25 minutes.

Consistent with previous literature (Dens et al., 2009), self-esteem and body esteem were significantly related ($r = .340$, $p < .001$). Moreover, male respondents reported significantly higher levels of both self-esteem ($M_{SE,M} = 5.50$, $sd = .98$) and body esteem ($M_{BES,M} = 3.90$, $sd = 1.43$) than the female participants in the study ($M_{SE,F} = 5.17$, $sd = 1.49$; $M_{BES,F} = 3.36$, $sd = 1.49$) (t_{SE} (201) = 2.11, $p = .036$; t_{BES} (203) = 2.54, $p = .012$). We carried out a median split on levels of self-esteem of the respondent (i.e., low (score 5.2 and lower) vs. high), and on levels of body esteem of the respondent (i.e., low (score 3.4 and lower) vs. high).

VI. RESULTS

A repeated-measures MANOVA, with body-size (i.e., relatively slim, average size, full-figured) of the female endorser as a within-subjects factor, revealed that there was a significant main effect on all the dependent measures of ad effectiveness (Aad, AB, PI) and the attitude towards the endorser (the attitude towards the inner characteristics as well as the attitude towards the outer characteristics of the female endorsers). Post-hoc tests show that the average scores for the relatively slim body-size ('ideal') model in the advertisement, for all the ad effectiveness measures, are significantly higher than for the full-figured model (at a significant level of 0.05). Moreover, the average-sized model scores significantly better than the full-figured model. This pattern is found for all metrics we evaluated at a 0.05 significant level. However, the level of fit between the respondent and the model is the highest for the average-sized model (we only analyzed this for the female participants, as we only used female endorsers in the advertisements). In Table 1 the average scores for the different dependent variables and the F-test results are presented.

Table 2. Average scores for the different attitudes towards the different print-ads of female endorsers*

	Aad	**AB**	**PI**	**Att-endorser-inner**	**Att-endorser-outer**	**fit (female) resp-endorser**
Slim body-size ('ideal')	**3.89 (1.58)**	**3.93 (1.17)**	**3.28 (1.17)**	**3.94 (1.15)**	**4.50 (1.47)**	2.13 (1.16)
Average body-size	3.56 (1.59)	3.64 (1.31)	3.07 (1.23)	3.79 (1.29)	4.08 (1.72)	**2.17 (1.33)**
Full-figured body-size	2.81 (1.31)	3.26 (1.17)	2.58 (1.02)	3.52 (1.29)	2.99 (1.52)	1.75 (1.06)
F-value df (p-value)	24.55 [2, 362] ($p < 0.001$)	17.07 [2, 362] ($p < 0.001$)	19.38 [2, 362] ($p < 0.001$)	5.09 [2, 362] ($p = 0.007$)	45.95 [2, 362] ($p < 0.001$)	7.30 [2, 252] ($p = 0.001$)

* Post-hoc tests show significant differences between the different groups (slim, average, full-size) for Aad, AB, PI, Att-endorser-inner and Att-endorser-outer.

To answer the third research question related to product characteristics and individual differences of the participants we introduced different moderator variables in the original MANOVA model. First, we found no significant main effect (and or interaction effect) for the product characteristics. Second, we took socio-demographic characteristics, gender on the one hand and age on the other hand in the original MANOVA model and found that overall, men and women of different age groups did not respond in significantly different ways to the

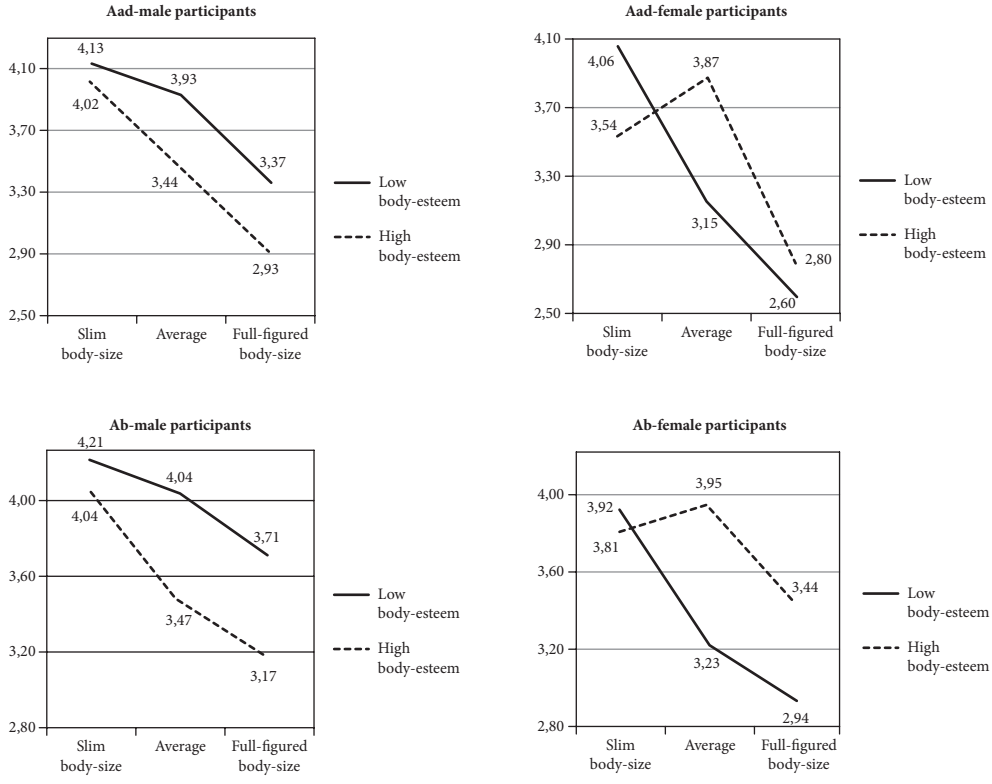
different endorsers in the ads. However, only one significant difference was found for gender for the ad with the full-figured model, where the average scores of the male audience were significantly higher for the Aad ($M_{\text{Aad, female resp, full-figured body-size}} = 2{,}71$, $M_{\text{Aad, male resp, full-figured body-size}} = 3{,}11$; $t_{\text{gender}}(194) = 2.08$; $p = 0.039$). Thirdly, we also included psychographic characteristics in the original MANOVA model. Again, neither variable, self- and body-esteem, significantly influenced the pattern of the responses to the different ads as is seen in Table 1.

Finally, we ran two different MANOVAs by taking into account both the socio-demographic and the psychographic characteristics of the participants. In doing this, the two psychographic factors were individually entered as a third factor in the original 2 (gender of participant) x 3 (body-size of the endorser used in the ad) MANOVA. With self-esteem (high versus low), we found no significant interaction effects on the different dependent measures. However, with respect to body esteem, we did find a significant three-way interaction effect on Aad ($F[2, 344] = 3.564$, $p = .029$) and AB ($F[2, 344] = 3.454$, $p = .033$). Judging from the average Aad scores of the male respondents with low and high levels of body-esteem, both groups preferred the relatively slim model ($M\text{-}_{lowbodyesteem} = 4.13$, $sd = 1.57$; $M_{\text{-}high} = 3.98$, $sd = 1.64$) over the average model ($M\text{-}_{low} = 3.93$, $sd = 1.49$; $M_{\text{-}high} = 3.51$, $sd = 1.64$) and the full-figured model ($M\text{-}_{low} = 3.37$, $sd = 1.30$; $M_{\text{-}high} = 2.94$, $sd = 1.31$). Also, female participants with a low score on body esteem have significantly higher scores for the relatively slim (ideal) female endorsers. However, female participants with a high score for body esteem show a significantly higher score for the average body-size female endorsers. The average Aad scores of the female (low – and high body esteem) participants for the ad with the image of the slim body-size endorser are (1) $M\text{-}_{low} = 4.06$ ($sd = 1.44$) and $M_{\text{-}high} = 3.56$ ($sd = 1.64$), (2) for the average body-size endorser $M\text{-}_{low} = 3.31$ ($sd = 1.50$) and $M_{\text{-}high} = 3.84$ ($sd = 1.64$) and (3) for the full-figured body-size endorser $M\text{-}_{low} = 2.59$ ($sd = 1.38$) and $M_{\text{-}high} = 2.80$ ($sd = 1.28$).

On Ab, we find the same pattern, Both high and low body esteem male participants prefer the slim body-size female endorser over the other models ($M\text{-}_{low\text{-}slim} = 4.21$, $sd = 1.15$; $M_{\text{-}high\text{-}slim} = 3.98$, $sd = 1.21$; $M\text{-}_{low\text{-}average} = 4.04$, $sd = 1.11$; $M_{\text{-}high\text{-}average} = 3.47$, $sd = 1.50$; $M\text{-}_{low\text{-}full\text{-}figured} = 3.71$, $sd = 1.02$; $M_{\text{-}high\text{-}full\text{-}figured} = 3.19$, $sd = 1.25$). However, the female participants with a low level of body esteem have a significantly higher score for Ab after viewing the ad with the slim body-size female endorser, whereas the female participants with a high score for body esteem have a significantly higher score for the ad with the average body-size female endorser ($M\text{-}_{low\text{-}slim} = 3.89$, $sd = 1.21$; $M_{\text{-}high\text{-}slim} = 3.82$, $sd = 1.14$; $M\text{-}_{low\text{-}average} = 3.36$, $sd = 1.33$; $M_{\text{-}high\text{-}average} = 3.95$, $sd = 1.10$; $M\text{-}_{low\text{-}full\text{-}figured} = 2.97$, $sd = 1.16$; $M_{\text{-}high\text{-}full\text{-}figured} = 3.48$, $sd = 1.11$). For PI we have found the same pattern, however, the average scores did not differ significantly. Our results show that levels of body esteem and not self-esteem significantly influence the ad effectiveness. This result is consistent with earlier findings of Dens et al. (2009), where the ad effectiveness of highly versus less scarcely dressed female endorsers was examined and body esteem and not self-esteem plays a significant role. Dens et al. (2009) suggested that self-esteem is a more general and robust trait, whereas body esteem (through comparison i.e. self-evaluation) can be seen as more specific.

In Figure 5, these three-way interaction effects are visualized for Aad and AB for the male and the female participants separately. The figure also displays women and men's Aad and AB mean scores for body esteem for the different conditions We can see that the female participants with relatively high levels of body esteem gave the highest Aad and AB scores for the image with the average-size model. For females with a relatively low score for body esteem, the highest score goes to the slim 'ideal' female endorser and the gap between the slim ('ideal') body-size and the other endorsers is significantly different as compared to the other average body-size and full-figured body-size endorsers. For the male participants, we did not find these significant differences, that is to say levels of body esteem did not significantly influence their responses to differently shaped body-size female endorsers. Male respondents always preferred the slim ('ideal') body-size endorser over the other endorsers.

Figure 5. Three way interaction degree of body esteem of participant by gender of participant on difference in ad effectiveness (Aad and AB)



VII. STUDY 3

A. DESIGN AND PROCEDURE

In the third study, we analysed the ad effectiveness of the different body-sizes of male endorsers. As we found no significant differences between the different beauty products in the first study, we decided to use different (body-size related) products in the advertisements of study 3. We opt for food products and more specifically distinguish between unhealthy and healthy food products as a body-size related product in the study. The literature review also shows that exposing previewers to idealized (body-size) endorsers could trigger eating disorders (Halliwell, Dittmar & Howe, 2005; Bessenoff, 2006). This suggests that body-size as a moderating factor may be particularly effective in advertisements of food related products. Furthermore, the product match-up model (Bower and Landeth, 2001) suggests that endorsers with an 'ideal' body-size can be relatively more effective in endorsing healthy products than their counterparts.

B. PRETEST SELECTION (UN)HEALTHY PRODUCTS

The participants in the pretest were exposed to 10 different products (i.e., an apple, French fries, whole wheat bread, salad, hamburger, cereals, water, chocolate bar, spaghetti bolognaise and fresh orange juice). The products were tested for their 'health' image, on a 5-point-Likert scale ('*Please indicate how healthy you find this product where 1 is very unhealthy and 5 is very healthy*'). In total 20 different participants, via a self-selected online survey, completed the questionnaire. The results of the repeated measures ANOVA indicated significant differences between the products ($F[9,135] = 85{,}59$; $p < .001$). The hamburger was found the most unhealthy product ($M = 1{,}00$; $sd = 0{,}00$), the apple the healthiest ($M = 4.94$; $sd = 0.25$) and the spaghetti bolognaise neither healthy nor unhealthy ($M = 3.25$; $sd = 0.86$). Based on the results of this pre-test, we chose an unhealthy product (the hamburger), a neutral product (the spaghetti) and a healthy product (the apple). In the fictitious advertisements, the products all received fictitious brand names to control for existing brand preferences and knowledge which could possibly distort respondents' evaluation of the advertisements.

The three selected products were used for the advertisements with the different body-size male endorsers. Again we investigated three unknown, professional endorsers who differed in their general looks, but now focused on male endorsers only. As is discussed in the introduction and Study 1, most men aspire to a muscular mesomorph shape, which is characterized by well-developed muscles on chest, arms and shoulders (Grogan, 2008). Grogan (2008) even indicates that men who are dissatisfied with their body shape would like to have more muscles (e.g. biceps, shoulders, chest) which can actually mean both an increase or a reduction in body weight, unlike 'dissatisfied' women who mostly want to

become slimmer. In Study 3, the average measurements of the 'average' body-size of the endorser was European shirt size = medium, chest = 100 cm, waist = 86 cm and hips = 105 cm. These sizes are comparable to those of the 'average' Belgian male fashion model (as listed on the agency websites, www.models.be). The slim model, in our study, has average European shirt size small (chest = 80 cm; waist = 70 cm; hips = 86 cm) and the full-figured / muscular ('ideal') endorser has average European shirt size large (chest = 112 cm, waist = 98 cm, hips = 116 cm).

Again, a 3 × 3 Latin square mixed factorial design was set up for the different body-size male endorsers for the three different food products ranging from unhealthy, neither unhealthy nor healthy, to healthy with fictitious brand names.

C. SAMPLE

As in study 2, male and female participants were recruited to fill in the online questionnaire. Participants were asked to evaluate three advertisements. All respondents saw all three specifically healthy, neutral and unhealthy products and models of all three body-sizes, but in a different order and in different product body shape combinations. The main data collection was carried out via a self-selected online survey of master degree students and their relatives at an urban university in Belgium. Incomplete responses and participants with the same responses for all questions were deleted. In total, 573 fully completed questionnaires were obtained. In this sample, 64% of the participants were female and 36% were male. 88.5% of the participants were between 18 and 25 years. No significant differences were found for age and gender between the different experimental groups ($X^2_{gender} = 6.82$, $p = .56$; $X^2_{age} = 5.86$, $p = .66$).

D. RESULTS

The different Cronbach's alpha scores and the results of the explorative factor analyses indicate that also for Study 3 (identical to Study 2) we were allowed to construct Aad, AB, PI, attitude towards the inner- and outer characteristics of the endorser, the fit between the endorser and the (male) participant, self- and body esteem as the concepts for our dependent measures. As in Study 2, we carried out a median split on the level of self-esteem (i.e., low (score 4.8 and lower) vs. high), and body esteem of the participant (i.e., low (score 4.5 and lower) vs. high).

A repeated measures MANOVA, with a body-size of the male endorser and product as within-subjects factors, revealed that there was a significant main effect of the body-size of male endorser on the dependent measures and (again) no significant main effect for the products. Study 3 indicates that for the male models, the 'ideal' body-size male endorser, i.e. the muscular male model, and the average body-size endorser score significant higher for the

ad effectiveness measurements compared to the slim body-size male endorser. The post-hoc tests show significant higher scores for the 'ideal' muscular – and average body-size 'male' endorsers for PI, Aad, the attitude towards the out-side characteristics of the endorser and the fit between endorser and participant compared to the slim body-size endorser. No significant differences could be found for the attitude towards the brand/product (Ab) and the different body-sizes of the male endorser in the advertisement. Moreover, the score of the attitude towards the inner characteristics of the endorser are significantly higher for the average body-sized model compared to the others. Also in study 3, the level of fit between the participant and the endorser is significantly lower for the 'non-ideal' endorser when compared to the other models (this was only analysed for the male participants, as we only used male endorsers in the ad). This result is in line with earlier findings indicating that men tend to 'overestimate' their body-size, a tendency which often means that men believe that their body corresponds with that of muscular male 'ideal' bodies (Grogan, 2008).

In Table 3 the averages scores for the different dependent variables and the F-test results are presented.

Table 3. Average scores for the different attitudes towards the different print-ads of male endorsers (n = 573)*

	Aad	**AB**	**PI**	**Att-endorser-inner**	**Att-endorser-outer**	**fit (male) resp-endorser**
Slim body-size	2.62 (1.31)	3.48 (1.40)	2.96 (1.38)	2.91 (1.36)	3.27 (1.30)	1.95 (1.15)
Average body-size	**3.10 (1.39)**	3.49 (1.32)	**3.16 (1.40)**	**4.18 (1.36)**	**3.52 (1.20)**	**2.28 (1.20)**
Muscular ('ideal')	**3.09 (1.60)**	3.53 (1.51)	**3.30 (1.49)**	3.71 (1.47)	**3.46 (1.30)**	**2.31 (1.31)**
F-value df (p-value)	37.90 (2, 1144) ($p < 0.001$)	0.29 (2, 1144) ($p = 0.749$)	12.78 (2, 1144) ($p < 0.001$)	187.29 (2, 1144) ($p = 0.007$)	10.24 (2, 1144) ($p < 0.001$)	8.60 (2, 410) ($p < 0.001$)

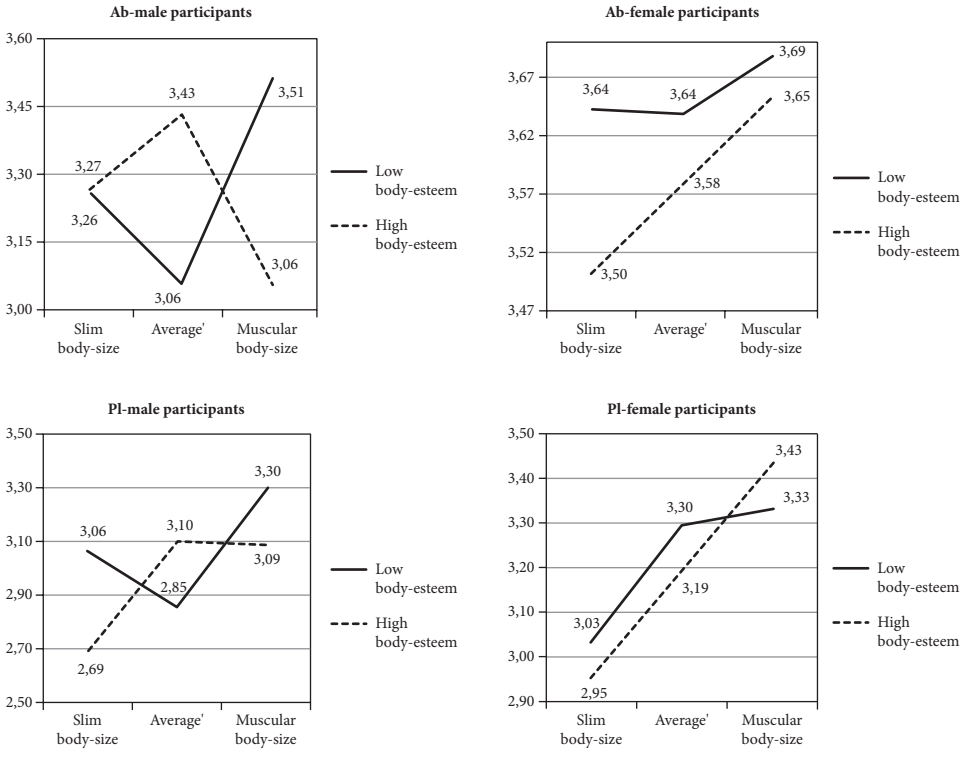
* Post-hoc tests show significant differences between the slim and average / muscular groups for Aad, PI, Att-endorser-inner, Att-endorser-outer.

The results for the 'ideal' male endorsers are less 'outspoken' than for the 'ideal' female endorsers. However, they are in line with that for the female endorsers where the ideal model (e.g. the relatively slim model) is the most effective. The literature indicates that the body of a non-muscular male model (in our study the slimmer male model) is the one that is most incongruent with an ideal model. Also the research results of Table 3 indicate that the slimmer body-size male models are less effective for print advertisements.

To answer the third research question for the male endorsers, we introduced different moderator variables in the original MANOVA model. As in Study 2, we found that men and women did respond differently in a significant way to the different body-size endorsers in the ads. However, for all the effectiveness measures, the female respondents score significantly higher than the male respondents. This suggests that the ad effectiveness is significantly higher for the female audience with opposite-sex models. Also in Study 3, age does not significantly influence the research results. Furthermore, we also included self- and body esteem in the original MANOVA model. Consistent with study 2, self-esteem and body esteem were significantly related (r = .251, p < .001) and also in this sample men reported significantly higher levels of both self-esteem ($M_{SE,M}$ = 5.22, sd = .93) and body esteem ($M_{BES,M}$ = 5.07, sd = 1.50) than women ($M_{SE,F}$ = 4.77, sd = 0.89; $M_{BES,F}$ = 4.56, sd = 1.45) (t_{SE} (571) = 5.75, p < 0.001; t_{BES} (571) = 5.07, p < 0.001). Finally, we ran two different MANOVAs by taking into account both gender and the psychographic characteristics of the participants. To do this, the two psychographic factors were entered as a third factor in the original 2 (gender of participant) × 3 (male model used in the ad) MANOVA. With self-esteem (high versus low), we found in this study, as in the others, no significant interaction effects on the different dependent measures. But we found again a significant three-way interaction effect with respect to body esteem.

As in Study 2, we found this effect on Ab (F[2, 1138] = 4.245, p = .015) and also on PI (F[2, 1138] = 3.319, p = .037), but not on Aad. Looking at the average Ab scores of the male respondents with low and high levels of body esteem, we can conclude that male respondents with relatively high body esteem prefer the 'average' body-size endorser over the 'ideal' muscular and the relatively slim model ($M\text{-}_{high\text{-}body\text{-}esteem,\ avg\ endorser}$ = 3.43, sd = 1.36). The male respondents with relatively low body esteem score preferred the muscular ('ideal') endorser ($M\text{-}_{low\text{-}body\text{-}esteem,\ muscular\text{-}endorser}$ = 3.51, sd = 1.58). The female participants with low and high levels of body esteem score the muscular ('ideal') endorser the highest for Ab and PI. In Figure 4, these three-way interaction effects are visualized for Ab and PI for the male and the female participants separately. We can see that the male participants with relatively high body esteem have the highest Ab and PI scores for the image with the average body-size model. For males with a relatively low score on body esteem the highest score is for the muscular ('ideal') male endorser. For the female participants, we did not find significant differences, that is to say body esteem did not significantly influence their responses to differently shaped male body-size endorsers. Female respondents always preferred the muscular ('ideal') endorser over the other models. While the effects are not as pronounced as in Study 2, overall, Study 3 represents the results of Study 2 for male endorsers.

Figure 6. Three way interaction degree of body esteem of participant by gender of participant on difference in ad effectiveness (AB and PI)



VIII. DISCUSSION

On the basis of our research results, we conclude that an advertiser selecting a female non-celebrity endorser for a print advertisement to promote a beauty product should generally choose a slim model (size 34). These models are significantly more effective than the average-size or full-figured endorsers. This is valid for both male and female audiences (and for all age groups).

The research results for male non-celebrity endorsers suggest that print ads should rely on a full-figured muscular model and not on slim male endorsers. The full-figured muscular male model has in almost all cases significantly higher scores on advertising effectiveness measurements as compared to the slim male models, especially in the eyes of a female audience. These results suggest that female and male "ideal type" endorsers – slim and muscular, respectively – are generally the most effective in advertisements. This result suggests that despite the feelings of unease which ideal body sized endorsers might generate

in audiences, their use nevertheless leads to the greatest ad effectiveness. However, our analysis further suggests that body-esteem is an important and moderating variable, in particular for women. That is to say, women with a relatively low score on body esteem find the slim body-size endorsers used in the advertisement the most effective, whereas women with a relatively high score for body esteem find the medium body-sized endorsers more effective. This suggests that, as a general rule, slim body-size female endorsers should be used, unless the advertisement targets women with a high level of body esteem, in which case medium body-sized endorsers should be considered. A more differentiated approach in promoting beauty products to women could therefore prove to be useful. The well-known Dove campaign, where beauty products of Dove are advertised with relatively medium body-size/ full-figured female models, could therefore have been effective as a form of product differentiation targeted at a sub-set of women.

This result could be explained by findings in earlier research by Mathes and Kahn (1975) and Rosenberg (1965), which suggests that the thin ideal is strongly related to the overall self-esteem of women. Women with high body esteem are, therefore, unlikely to aspire to a different "ideal" body size and will not identify with endorsers with ideal body sizes but more easily with "normal" or "average" body sizes (assuming that on average they have average body sizes). The identification effect (Wilcox & Laird, 2000), which generally leads to positive feelings and a more positive attitude towards the ad, will, therefore, occur more strongly with regard to endorsers with non-ideal body sizes which, in turn, would enhance the ad effectiveness for these endorsers.

Our results indicate the same heterogeneity in the responses of the male audience to the male endorsers, although less marked. Also for a male audience with high levels of body esteem the use of non-ideal body-size male models might be more effective (a "male" Dove campaign could, therefore, be considered).

It should also be noted that the impact on ad effectiveness of using models of different body sizes differs significantly between male and female models. More specifically, the research results indicate that the slimmer (ideal) body-size female models – except for the high body esteem female audience – are significantly more effective. For male endorsers the impact of using an ideal body-size model is not as pronounced in the advertising effectiveness results. These results are independent of the gender of the audience. This could suggest that for male endorsers body-size does not play such an important role in defining the 'ideal' male model and/or that other factors influence the way ideal male endorsers are portrayed for print advertisements. It is also possible that an 'ideal' male model is less pronounced as compared to female models and this is perhaps due to the fact that male endorsers are still not that familiar in the ad industry when compared to the more frequently used female endorsers. However, this might be changing through new marketing initiatives[1]

[1] For example, the recent marketing by Abercrombie & Fitch emphasizes the use of full-figured muscular male models (www.youtube.com/watch?v=z5NRWM3FgqA).

Table 4 shows the impact on attitudes towards the brand from using an 'ideal' endorser in a print advertisement. The results show that the impact on effectiveness when using the 'ideal' female endorser is very large for the male audience (see Table 4, for the male audience with low BE: +13.48%, high BE +24.76%) and even larger for the female audience with a low score on body esteem (+30.98%). Using an 'ideal' body-size male endorser for a male audience with a high score for BE will decrease the AB scores by, respectively, 10.79% and 3.29% for using an 'ideal' body-size female endorser for a female audience as compared to using an average body-size endorser. The influence of using 'ideal' body-size male models is – when compared to the female endorsers – much smaller. For example, in the case of the female audience with low BE, an improvement of only +1.37% is found and for the female audience with high BE +2.00% (see Table 4). The AB measurement of the print ads increases by +14.71% for the male audience when an ideal body-size male endorser is used as compared to the average body-size endorser.

Table 4. Significant advertising effectiveness improvements in percentage of using ideal model(1) compared to "average body-size" endorser (for same and opposite sex models, Study 2 and 3)

Participant	Male		Female	
	Low-BE	High-BE	Low-BE	High-BE
AB-(using ideal opposite sex model)	+13.48%	+24.76%	+1.37%	+2.00%
AB-(using ideal same sex model)	+14.71%	-10.79%(2)	+30.98%	-3.29%(2)

(1) Ideal female model: slim body-size endorser; ideal male model: muscular model

(2) Improvement when using ideal model according to definition (1) instead of using the average body-size model (ideal model for the high body esteem participants).

Table 4 indicates that an 'ideal' body-size female model has a much larger impact on a male audience than an 'ideal' body-size male model has on a female audience. Furthermore, Table 4 also suggests that the audience (male and female) with a high score on body esteem is relatively more susceptible to an 'ideal' body-size endorser from the opposite sex than are their peers with low body esteem.

Based on our findings further research should include more analysis of ideal male body-size endorsers for advertising purposes. Furthermore, it would be interesting to analyse to what extent generational and cultural differences affect the relative effectiveness of the body sizes of endorsers. Extending the research to other countries (our research results are obtained from a sample of the Belgian population) would be necessary in order to obtain results that can be more easily generalised. The analysis would therefore have to be replicated in different countries and across different population segments to assess whether the results can be generalised.

This paper is based on a limited number of products; further research could usefully extend this to other product categories. Finally, this analysis has limited itself to unknown and relatively young endorsers, leaving open the question of whether older endorsers and/or celebrities are perceived differently for some product categories. This again, should be addressed in follow-up research.

REFERENCES

Agliata, D. & Tantleff-Dunn, S. (2004). The impact of media exposure on males' body image. *Journal of Social and Clinical Psychology*, 23(1), 7.

Bessenoff, G.R. (2006). Can the media affect us? Social comparison, self-discrepancy, and the thin ideal. *Psychology of Women Quarterly*, 30 (3), 239–251.

Bissell, K., & Rask, A. (2010), Real women on real beauty: self-discrepancy, internalization of the thin ideal, and perceptions of attractiveness and thinness in Dove's campaign for real beauty, *International Journal of advertising, 29* (4), 643–668.

Bottamini, G., & Ste-Marie, D. (2006), Male Voices on Body Image, *International Journal of Men's Health*, 5 (2), 109–132.

Bower, A.B. (2001). Highly attractive models in advertising and the women who loathe them: the implications of negative effect for spokesperson effectiveness. *Journal of Advertising*, 30 (3), 51–63.

Bower, A.B., & Landreth, S. (2001). Is beauty best? Highly versus Normally attractive models in advertising. *Journal of advertising*, 30 (1), 1–12.

Costa, M. (2011). Interest rate in male beauty brands rises, *Marketingweek*, September, 15.

Dens, N., De Pelsmacker, P., & Janssens, W. (2009). Effects of scarcely dressed models in advertising on body esteem for Belgian men and women. *Sex Roles*, 60 (5/6), 366–378.

Diedrichs, P.C. & Lee, C. (2010), GI Joe or Average Joe? The impact of average-size and muscular male fashion models on men's and women's body image and advertisement effectiveness, *Body Image*, 7, 218–226.

Eisend, M., Langner, T. (2010) 'Immediate and delayed advertising effects of celebrity endorsers attractiveness and expertise', International Journal of Advertising, 29(4), 527–546.

Erdogan, B.Z. (1999) 'Celebrity endorsement: A literature review', Journal of Marketing

Erdogan, B.Z., Baker, M.J., & Tagg, S. (2001) 'Selecting celebrity endorsers: The practitioner's perspective', Journal of Advertising Research, 41(3), 39–48.

Gelb, B.D., & Zinkhan, G.M. (1986). Humor and advertising effectiveness after repeated exposure to a radio commercial. *Journal of Advertising, 15 (2),* 15–34.

Gielen, G. (2003). Onaantrekkelijk?: beeldvorming over het belang van fysieke aantrekking. Antwerpen: Garant.

Grabe, S., Ward, L.M., Hyde, J.S. (2008). The role of the media in body image concerns among women: A meta-analysis of experimental and correlational studies, *Psychological Bulletin*, 134(3), 460–476.

Groesz, L.M., Levine, M.P., & Murnen, S.K. (2002). The effect of experimental presentation of thin media images on body satisfaction: A meta-analytic review. *International Journal of Eating Disorders*, 31, 1–16.

Grogan, S. (2008), Body image: understanding body dissatisfaction in men, women, and children, Routledge, 250 p.

Gustafson, R., Popovich, M., & Thomsen, S. (1999). The thin ideal. *Marketing News*, 15, 22.

Halliwell, E., & Dittmar, H. (2004). Does size matter? The impact of model's body size on advertising effectiveness and women's body-focused anxiety and advertising effectiveness, *Journal of Social and Clinical Psychology*, 23 (1), 104–122.

Halliwell, E., Dittmar, H. & Howe, J. (2005), The Impact of Advertisements Featuring Ultra-thin or Average-size Models on Women With a History of Eating Disorders, *Journal of Community & Applied Social Psychology*, 15, 406–413.

Harrison, K. & Cantor, J. (1997). The Relationship between Media Consumption and Eating Disorders; *Journal of Communication*, 47, 40–67.

Henderson-King, E., & Henderson-King, D. (1997). Media effects on women's body esteem: Social and individual difference factors. *Journal of Applied Social Psychology, 27* (5), 399–417.

Higgins, E.T. & Brendl, C. M. (1995). Accessibility and applicability – Some activation rules influencing judgment. *Journal of Experimental Social Psychology*, 31(3), 218–243.

Hobza, C.L., Walker, K.E., Yakushko, O. & Peugh, J.L. (2007). What about Men? Social Comparison and the Effects of Media Images on Body and Self-Esteem, *Psychology of Men & Masculinity*, 8(3), 161–172.

Irving, L.M. (1990). Mirror images: Effects of the standard of beauty on the self- and body esteem of women exhibiting varying levels of bulimic symptoms. *Journal of Social and Clinical Psychology*, 9, 230–242.

Jalees, T. & Majid, H. (2009). Impact of 'Ideal Models' Being Portrayed by Media on Young Females, *Paradigm*, 13(1), 11–19.

Jamieson, L. (1989). Adjusting stated intention measures to predict trial purchase of new products: A comparison of models and methods. *Journal of Marketing Research*, 26(3), 336.

Joshi, R., Herman, C.P., & Polivy, J. (2004). Self-enhancing effects of exposure to thin-body images. *International Journal of Eating Disorders*, 35(3), 333–341.

Kang, Y.S. & Herr, P.M. (2006). Beauty and the beholder: Toward an integrative model of communication source effects. *Journal of Consumer Research*, 33(1), 123–130.

Kilbourne, W. (1990). Female Stereotyping in Advertising: An Experiment on Male-Female Perceptions of Leadership. *Journalism Quarterly,* 67 (1), 25–31.

Lorenzen, L.A., Grieve, F.G., & Thomas, A. (2004). Exposure to muscular male models decreases men's body satisfaction. *Sex Roles, 51*(11/12), 743–748.

Marketingheart.wordpress.com (2011), http://marketingheart.wordpress.com/2011/02/19/in-a-world-of-fake-beauty-authenticity-can-still-pack-a-punch.

Martin, B.A.S., Veer, E. & Pervan, S.J. (2007), Self-referencing and consumer evaluations of larger-sized female models: A weight locus of control perspective, *Marketing Letters*, 18(3), 197–209.

Mathes, E.W., & Kahn, A. (1975). Physical attractiveness, happiness, neuroticism and self-esteem. *The Journal of Psychology*, 90 (1), 27–30.

McCracken, G. (1989) 'Who is the celebrity endorser? Cultural foundations of the endorsement process', Journal of Consumer Research, 16, pp. 310–321.

Mendelson, B., McLaren, L., Gauvin, L. & Steiger, H. (2002). The relationship of Self-Esteem and Body esteem in women with and without eating disorders, *International Journal of Eating Disorders*, 31 (3), 318–323.

Mahoney, S. (2011) Nielsen: Men take to stores, women head to couch, *MarketingDaily*, 15–3–2011.

Milkie, M. A. (1999). Social comparisons, reflected appraisals, and mass media: The impact of pervasive beauty images on Black and white girls' self-concepts. *Social Psychology Quarterly, 62* (2), 190–210.

Milkie, M. A. (2002). Contested images of femininity. *Gender and society, 16* (6), 839–859.

Polivy, J., & Herman, C.P. (1987). Diagnosis and treatment of normal eating. *Journal of Consulting and Clinical Psychology*, 55 (5), 635–644.

Posavac, H.D., Posavac, S.S. & Posavac, E.J. (1998). Exposure to media images of female attractiveness and concern with body weight among young women. *Sex Roles, 38* (3/4), 187–201.

Roozen, Irene and Christel Claeys (2010), The Relative Effectiveness of Celebrity Endorsement for Print Advertisement, *Review of Business and Economic Literature*, Vol LV (January-March), 76–90.

Rosenberg, M. (1965). *Society and the adolescent self-image*. New Jersey: Princeton University Press.

Smeesters, D. Mussweiler, T., Mandel, N. (2010). The effects of thin and heavy media images on overweight and underweight consumers: Social comparison processes and behavioral implications. *Journal of Consumer Research*, 36(6), 930–949.

Smolak, L., Murnen, S.K. & Thompson, J.K. (2005) Sociocultural influences and Muscle Building in Adolescent Boys, *Psychology of Men & Masculinity*, 6, 227–239.

Solomon, M.R., Bamossy, G., Askegaard, S., Hogg, M.K., & Verhagen, W. (2008). Consumentengedrag. Amsterdam: Pearson Education.

Spears, N., & Singh, S. (2004). Measuring attitude toward the brand and purchase intentions, *Journal of Current Issues & Research in Advertising,* 26 (2), 53–66.

Till, B.D. & Busler, M. (2000). The match-up hypothesis: Physical attractiveness, expertise, and the role of fit on brand attitude, purchase intent and brand beliefs. *Journal of Advertising*, 29(3), 1–13.

Wilcox, K., & Laird, J.D. (2000). The Impact of Media Images of Super-Slender Women on Women's Self-Esteem: Identification, Social Comparison, and Self-Perception. *Journal of Research in Personality*, 34 (2), 278–286.